



AIVR 2026

The Third International Conference on Artificial Intelligence and Immersive Virtual
Reality

ISBN: 978-1-68558-381-1

April 19 - 23, 2026

Lisbon, Portugal

AIVR 2026 Editors

Muneo Kitajima, Nagaoka University of Technology (Emeritus), Japan

AIVR 2026

Forward

The Third International Conference on Artificial Intelligence and Immersive Virtual Reality (AIVR 2026), held on April 19 – 23, 2026, initiated a series of events addressing the interaction between Artificial Intelligence and Virtual Reality.

Industry, agriculture, finance, health, society, education and almost all domains, including human-systems interactions (interfaces, requests, trust, ethics, etc.) are subject of major evolution with the infusion with AI-based mechanisms into Virtual environments.

Virtual environments are deemed to shape the future society. Extended virtual world will be seamlessly integrated with the physical world creating digital twins. The convergence of computing, communication, and networking for supporting complex applications (huge data, complex processing algorithms) will benefit from the Artificial Intelligence (AI) progress, especially on Deep learning, Machine learning, and Data Analytics.

This event attracted excellent contributions and active participation from all over the world. We were very pleased to receive top quality contributions.

We take here the opportunity to warmly thank all the members of the AIVR 2026 technical program committee, as well as the numerous reviewers. The creation of a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to AIVR 2026. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the AIVR 2026 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope AIVR 2026 was a successful international forum for the exchange of ideas and results between academia and industry that will promote further progress in the area of AI and VR. We also hope that Lisbon provided a pleasant environment during the conference and everyone saved some time to enjoy this beautiful city.

AIVR 2026 Steering Committee

Muneo Kitajima, Nagaoka University of Technology (Emeritus), Japan
Carolina Cruz-Neira, University of Central Florida, USA
Aurelie Mailloux, Reims Hospital and University of Reims, France
Jerome Dinet, University of Lorraine, France
Zahra Moussavi, University of Manitoba, Canada

AIVR 2026

Committee

AIVR 2026 Steering Committee

Muneo Kitajima, Nagaoka University of Technology (Emeritus), Japan
Carolina Cruz-Neira, University of Central Florida, USA
Aurelie Mailloux, Reims Hospital and University of Reims, France
Jerome Dinet, University of Lorraine, France
Zahra Moussavi, University of Manitoba, Canada

AIVR 2026 Technical Program Committee

Suayb S. Arslan, Boğaziçi University, Turkey
Mahajan Sagar Bhaskar, Prince Sultan University (PSU), Saudi Arabia
Concettina Buccella, University of L'Aquila, Italy
Pasquale Cascarano, University of Bologna, Italy
Jean-Rémy Chardonnet, Institut Arts et Métiers de Chalon-sur-Saône | French-German Institute for Industry of the Future (ENSAM/KIT), France
Mathieu Chollet, University of Glasgow, UK
Courtney D. Cogburn, Columbia University, USA
Carolina Cruz-Neira, University of Central Florida, USA
A. Augusto de Sousa, Universidade do Porto, Portugal
Massimiliano di Luca, University of Birmingham, UK
Jerome Dinet, University of Lorraine, France
Sai Anvesh Durvasula, Parabole.ai, USA
Christos Gatzidis, Bournemouth University, UK
Paulo Gil, NOVA School of Science and Technology, Lisbon, Portugal
Yan Hu, Blekinge Institute of Technology (BTH), Sweden
Florentine Hüttel, University Medical Center of the Johannes Gutenberg-University Mainz, Germany
Magdalena Igras-Cybulska, Interactive crafts sp. z o.o., Poland
Yirui Jiang, Cranfield University, UK
Celia Kessassi, Marshall University, USA
Muneo Kitajima, Nagaoka University of Technology, Japan
Nadia Magnenat-Thalmann, University of Geneva/MIRALab, Switzerland
Aurelie Mailloux, University of Reims (URCA) / Reims Hospital / 2LPN, France
Vicky McArthur, Carleton University, Canada
Sonu Mehta, Microsoft Research, India
Zahra Moussavi, University of Manitoba, Canada
Wolfgang Mueller, University of Education Weingarten, Germany
Moloud Nasiri, Converse University, USA
Michael Nkereuwem, Federal University Dutse, Nigeria
Stefania Palmieri, Politecnico di Milano, Italy

Stamatis Papadakis, University of Crete, Greece
George Papagiannakis, University of Crete, Greece
Vatesh Pasrija, Meta - Seattle, USA
Gerald Penn, University of Toronto, Canada
Krzysztof Pietroszek, American University, Washington, USA
Ladislav Polák, Brno University of Technology, Czech Republic
Simant Prakoonwit, University of Southampton, UK
Yiming Qiu, University of Michigan, USA
Frank Steinicke, University of Hamburg, Germany
Ibrahim A. Tahmid, Virginia Tech, USA
Davide Tebaldi, University of Modena and Reggio Emilia, Italy
Yanbo Wang, Aalborg University, Denmark
Zixuan Wang, UC San Diego, USA

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission to reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Making VR Environments Fully Resonate with Interacting Humans: Core Constructs <i>Muneo Kitajima, Makoto Toyota, and Katsuko T. Nakahira</i>	1
C-LISTEN: Cognitive Load Impacts of Sensory-Triggered Environmental Navigation in Virtual Reality <i>Md. Monowar Hossain and M. Rasel Mahmud</i>	10
Human-Centered Innovation in Water Resources Management Education. A Co-Creation Approach to Developing VR Learning Experiences <i>Stella Kanatouri, Heribert Nacken, Sina Truckenbrodt, and Friederike Klan</i>	16

Making VR Environments Fully Resonate with Interacting Humans: Core Constructs

Muneo Kitajima 

Nagaoka University of Technology
Nagaoka, Niigata, Japan

Email: mkitajima@kjs.nagaokaut.ac.jp

Makoto Toyota

T-Method
Chiba, Japan

Email: pubmtoyota@mac.com

Katsuko T. Nakahira 

Nagaoka University of Technology
Nagaoka, Niigata, Japan

Email: katsuko@vos.nagaokaut.ac.jp

Abstract— Much of human life is shaped by interacting with artificial objects in the surrounding environment. When these artificial objects behave not according to natural laws but according to algorithms created by artifact designers, the environment can be considered a virtual environment. Our previous research has conceptualized the human-object interaction process as comprising perceptual, cognitive, and motor processes operating synchronously with the environment’s changes, alongside memory processes operating asynchronously with the environment’s changes, further linked by resonance between them. More specifically, P-resonance triggered by the basic senses comprised of rhythmic, spatial, and number senses connects perceptual process with the Perceptual-Multi-Dimensional Memory Frame, and C-resonance connects the Multi-Dimensional Memory Frame with cognitive process via the structure of Goals, Operators, Methods, and Selection rules (GOMS) established in the Multi-Dimensional Memory Frame. This paper presents the characteristics objects in a virtual environment should possess to enable smooth human-object interactions.

Keywords- Resonance; Basic Senses; Meme; GOMS; MHP/RT; Symbol Grounding.

I. INTRODUCTION

Humans acquire information about the external world through their five senses, activate Perceptual, Cognitive, and Motor (PCM) processes, and select and execute appropriate actions based on the circumstances at that moment. These processes are controlled by the brain. The brain forms a complex mechanism composed of an immense number of cells. However, when viewed functionally as an information processing device that appropriately converts input information to the brain and outputs it to the outside world, it can be understood as a remarkably simple cognitive architecture model.

Card, Moran, and Newell [1] devised Model Human Processor (MHP), a cognitive architecture model that simulates human perception, cognition, and motor processes. By analogy to von Neumann-type computers as information processing devices, they conceived the brain as comprising the components “memory,” “processor,” and “program,” which defines how memory and processor are coupled. Memory is characterized by its storage capacity, decay time, and data type, while a processor is characterized by its cycle time, which is the processing time per unit task. Memory consists of visual and auditory image

stores, working memory, and long-term memory, each serving distinct functions. The following situation was envisioned as the simulation target domain for operating the information processing device specified above. This situation involves solving problems within a clearly defined problem space based on the principle of rationality to achieve the specified goal. The knowledge required in this situation is organized as Goals, Operators, Methods, and Selection rules (GOMS).

By encoding perceived information within the perceptual process and representing it as symbols, it becomes possible to *think* within the cognitive process. In the thinking process, memory is utilized, and symbols are successively transformed. Memory itself is also expressed through symbols. Part of the result of thinking provides an expression of a sequence of actions executable in the motor process. Based on the idea that human intelligence can be captured by the thinking process of manipulating symbols, Newell proposed the Physical Symbol System as a theory of human intelligence [2]. This approach provided the foundation for Soar, a cognitive architecture that views thinking as problem-solving process involving manipulation of a problem space [3][4].

However, the real-world environment with which humans interact changes its state moment by moment based on the mechanisms inherent within the environment itself. Mechanisms can be either linear or nonlinear. In the latter case, it is fundamentally impossible to predict the temporal evolution of the state. When confronted with such real-world environments, it is difficult to fully capture the appropriate actions required in response to the environment’s state as a rational problem-solving process within a predefined, well-defined problem space. We have developed a cognitive architecture model—the Model Human Processor with Realtime Constraints (MHP/RT)—that can simulate not only behaviors modeled by navigation within well-defined problem spaces targeted by MHP and Soar, but also adaptive and flexible behaviors [5][6]. MHP/RT integrates models for understanding individual behavioral ecology and models for understanding collective behavioral ecology, incorporating the latest research findings from neuroscience and cognitive science into the conceptual framework of MHP.

MHP/RT consists of the following three elements: Perception, Cognition, and Motor. These form the “Perception-Cognition-

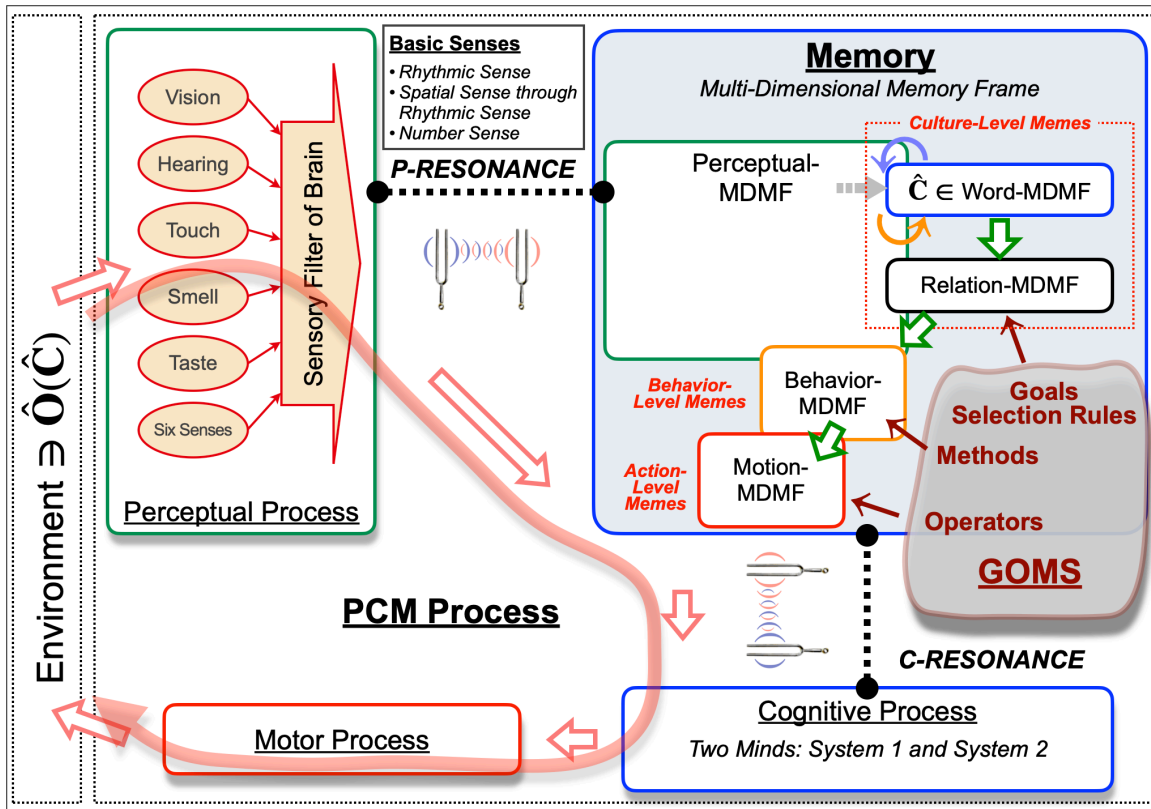


Figure 1. PCM processes operating synchronously with the environment and Multi-Dimensional Memory Frame operating asynchronously with the environment (created by synthesizing [7, Figure 2] and [8, Figure 6(b)] and making modifications).

Motor (PCM) Process,” which selects and executes actions that are likely to yield satisfactory results without causing the human system, as a living organism, to fail. This process *synchronizes* human interactions with the real-world environment along the time axis with the state transitions of that environment; the “Memory Process,” which operates *asynchronously* with the real environment, accumulating the results of PCM process execution and functioning to be utilized during action selection and execution; and a “resonance” mechanism connecting the environment-synchronized PCM process and the environment-asynchronous memory process. The operation of the PCM process represents individual behavioral ecology, while the contents of memory represent collective ecology—an accumulation of interactions with artifacts within a culture that is inherited across generations.

The key to understanding humans acting within their environment is as follows.

- (1) PCM processes operating in various modes depending on the task,
- (2) The structure and content of memory acquired while PCM processes are active and utilized by them,
- (3) P-resonance linking perceptual processes and memory,
- (4) C-resonance linking memory and cognitive processes,

Understanding is achieved by holistically relating these elements and grasping their overall dynamics.

Behavioral events occurring in actual dynamic environments

manifest in diverse forms. This paper identifies factors to consider when analyzing and understanding them, structuring insights gained during our research process according to the four points listed above. From the perspective of ensuring *smooth* interaction with the environment, it outlines the necessary requirements for achieving it.

This remainder of this paper is organized as follows. Section II explains the PCM process, memory process, and the P-resonance and C-resonance that connect them. Section III presents considerations for designing smooth interactions as conditions that induce resonance. Section IV summarizes this research.

II. PCM, MEMORY PROCESSES AND RESONANCE

This section provides an overview of the PCM process and memory process within MHP/RT, a cognitive architecture capable of simulating everyday behavioral selections [5][6][9] as the foundation.

A. Interaction between Human and Objects in the Environment

Interaction between humans and the object $\hat{O} \in \text{Environment}$ in the environment is realized through PCM processes operated by humans, memory processes utilized by PCM processes and updated as a result of PCM process execution, and processes occurring within \hat{O} . Figure 1 shows its overall outline.

\hat{O} changes state over time. These changes may arise either from human actions affecting \hat{O} or from mechanisms inherent within \hat{O} itself. In either case, for humans to smoothly continue their interaction with \hat{O} , it is necessary to synchronize their own PCM process with changes in the state of \hat{O} .

Meanwhile, memory defines the content executed by the PCM process. Memory accumulates the results of the PCM process's execution. Furthermore, memory is provided during the execution of the PCM process. The processes of accumulation and provision are executed asynchronously with respect to the environment's temporal development.

PCM processes operating synchronously with the environment and memory processes operating asynchronously with the environment collaborate to realize human behavior within the environment—what is executed and how—by linking them through resonance. Resonance includes P-resonance occurring between perceptual processes and memory, and C-resonance occurring between cognitive processes and memory. These enable the PCM process to utilize memory and generate behavior within the environment.

B. PCM Process Operating Synchronously with the Environment

1) *PCM Process*: When interacting with objects in the environment, humans receive physical and chemical stimuli through sensory nerves located at the interface with the environment and take in environmental information in the body. Further, it generates bodily movements that are suitable for the current environment. The stable and sustainable relationship between the environment and the self is established through continuous coordination between the activity of the self and the resultant changes in the environment, which should affect the self's next action.

Figure 1—created by combining [7, Figure 2] and [8, Figure 6(b)] and making modifications—illustrates the process by which environmental information is incorporated into the body via sensory nerves, undergoes processing within the brain, and then acts upon the external world via motor nerves, based on MHP/RT [5][6]. This process involves memory, modeled as Multi-Dimensional Memory Frame shown in the upper right of Figure 1, and the PCM processes indicated by the red arrows in Figure 1. The Multi-Dimensional Memory Frame comprises the Perceptual-, Behavior-, Motor-, Relation-, and Word-Multi-Dimensional Memory Frame. The Perceptual-Multi-Dimensional Memory Frame overlaps with the Behavior-, Relation-, and Word-Multi-Dimensional Memory Frame. This allows activity to propagate from Perceptual- to Motor-Multi-Dimensional Memory Frame.

This section explains the relationship between the PCM process and the memory process. Perceptual information taken in from the environment through sensory organs *resonates* with information in the Multi-Dimensional Memory Frame, which is called P-Resonance [10]. In Figure 1, this process is indicated by the symbol $\bullet\text{---}\bullet$. Resonance occurs first in the Perceptual-Multi-Dimensional Memory Frame and activates the memory network. After that, the activation spreads to the memory

Synchronous Modes

- Mode 1: Unconscious mechanism driven mode
A single set of perceptual stimuli initiates feedforward processes at the BIOLOGICAL and COGNITIVE bands to act with occasional feedback from an upper band, i.e., COGNITIVE, RATIONAL, or SOCIAL.

- Mode 2: Conscious mechanism driven mode
A single set of perceptual stimuli initiates a feedback process at the COGNITIVE band, and upon completion of the conscious action selection, the unconscious automatic feedforward process is activated at the BIOLOGICAL and COGNITIVE bands for action.

Asynchronous Modes

- Mode 3: In-phase autonomous activity mode
A set of perceptual stimuli initiates feedforward processes at the BIOLOGICAL and COGNITIVE bands with one and another intertwined occasional feedback process from an upper band, i.e., COGNITIVE, RATIONAL, or SOCIAL.

- Mode 4: Heterophasic autonomous activity mode
Multiple threads of perceptual stimuli initiate respective feedforward processes at the BIOLOGICAL and COGNITIVE bands, some with no feedback and others with feedback from the upper bands, i.e., COGNITIVE, RATIONAL, or SOCIAL.

Figure 2. Four operation modes of MHP/RT and their relationship with the four bands in the time scale of Newell's human action [3, Figure 3-3].

networks that overlap with the Perceptual-Multi-Dimensional Memory Frame, and finally to the Motor-Multi-Dimensional Memory Frame. In cognitive processing based on the Two Minds [0], conscious processing (System 2) and unconscious processing (System 1) operate in an interrelated manner [8][10]. System 2 utilizes the Word- and Relation-Multi-Dimensional Memory Frame via C-Resonance, while System 1 draws on the Behavior- and Motor-Multi-Dimensional Memory Frame via the same mechanism. Motor sequences are then expressed according to the Motor-Multi-Dimensional Memory Frame. The memories involved in the production of actions are updated to reflect the traces of their use process and influence the future action selection process.

2) *Four Operation Modes*: Humans interact with the external environment and select appropriate actions to achieve behavioral goals through a cycle of PCM processes. In MHP/RT, the action selection process is controlled by System 1 and System 2. These systems cooperate to link perception and movement, and the degree of cooperation depends on the state of the external environment with which the MHP/RT interacts. Figure 2 shows the Four Operation Modes characterized by the relationship between System 1 and System 2. There are synchronous and asynchronous modes.

The interaction is carried out in one of the four operation modes. Interactions executed in Mode 1 or Mode 2 are sound in the sense that System 2 monitors whether actions executed by System 1 in a timely manner have deviated from the desired trajectory. If signs of undesirable events occurring are detected,

feedback control via System 2 activates, enabling a return to normal operation. Meanwhile, Mode 3 and Mode 4 contain unmonitored feedforward System 1 processes. This indicates that instability may be introduced into the interactions.

In Mode 2, System 2 frequently intervenes the PCM processes conducted by System 1. More precisely, the pace of interaction is controlled by System 2. The role of System 1 is to carry out the necessary PCM processes, to advance the main System 2–environment interactions. As interactions progress, whenever the situation changes, the content of actions is determined starting from System 2 and executed by System 1. This does not correspond to the situation described in this paper where smooth interaction is achieved.

In contrast, in Mode 1, System 2's intervention is weak. The external environment supports the automatic processes performed unconsciously by System 1, enabling smooth interaction. This is the smooth interaction addressed in this paper. In this study, we propose that this is achieved through the occurrence of P-resonance and C-resonance. Our previous paper [13] provides guidelines for designing artifacts that realize interactions for Mode 1 and Mode 2. The guideline for supporting Mode 1 interaction is as follows.

Guideline [B]

1. For a normal Mode 1 interaction, provide information to both System 1 and System 2, so that System 1-led processes can run smoothly.
2. For an intensive Mode 1 interaction, e.g., video games and e-sports, focus on System 1 support.

This study details the considerations for implementing the items in this guideline.

C. Memory Processes Operating Asynchronously with the Environment

When the PCM process is running, the contents of Perceptual-Multi-Dimensional Memory Frame are updated in response to the perceptual process, those of Word-, Relation-, and Behavior-Multi-Dimensional Memory Frame are updated in response to the cognitive process, and those of Motor-Multi-Dimensional Memory Frame are updated in response to the motor process. In the memory shown in the upper right of Figure 1, focusing on memories that serve as traces of PCM process operation, Multi-Dimensional Memory Frame is classified into Perceptual-, Word-, Relation-, Behavior-, and Motor-Multi-Dimensional Memory Frame. This represents a structured approach focused on *memory updates* associated with the execution of the PCM process.

Imitation—doing what one observes—forms the basis of behavior. Therefore, the results of imitative behavior are structured and represented within the Multi-Dimensional Memory Frame, and imitative behavior arises through the propagation of activity within the Multi-Dimensional Memory Frame. Imitation occurs between individuals, and imitative behaviors exist that are passed down across generations. Therefore, the Multi-Dimensional Memory Frame utilized by the PCM process

and updated through its execution can be organized from the perspective of *memes* inherited across generations [14].

Word is considered the archetype of meme [15]. Words, i.e., symbols, exist within the individual's Word-Multi-Dimensional Memory Frame as language. Through the use of language, individual languages or cultural languages are organized within the Multi-Dimensional Memory Frame. Individual language (behavioral-level memes) is the language used in person-to-person communication, encompassing not only direct usage but also potential metaphorical usage, as well as thesauri—lists of words grouped by synonyms or related concepts. Cultural language (cultural-level memes) is language used within a cultural context where a proper understanding of the established common sense within a specific community is essential for successful communication. These languages develop during the early developmental stage from birth to age three, associated with objects in the environment (motor-level memes) encoded within neural networks. These circulate among people and persist across generations [16].

The relationship between the three levels of memes mentioned above and Multi-Dimensional Memory Frame is as follows:

- C-memes represent culture stored in the Relation- and Word-Multi-Dimensional Memory Frame.
- B-memes represent behaviors in the environment stored in the Behavior-Multi-Dimensional Memory Frame.
- A-memes represent bodily actions stored in the Motor-Multi-Dimensional Memory Frame.

D. P-Resonance

This section explains the mechanism by which P-resonance occurs between the perceptual process and Perceptual-Multi-Dimensional Memory Frame, based on our previous work [10].

1) *Characteristics of the Perceptual-Multi-Dimensional Memory Frame*: Humans act within an environment that changes moment by moment. The actions generated are observed as a continuous sequence of events containing environmental information in time and space.

Memory is formed by sampling this continuous sequence. Therefore, at the instance when memory is formed, information regarding absolute time and spatial coordinates is lost. The memory represented by the Multi-Dimensional Memory Frame, which has no position and time data, represents relationships between objects. Therefore, the contents of the memory can be reused by performing a topological transformation to match the current time and spatial scales. Here, by binding spatiotemporal information about actions to memories that contain only relational information between events, these memories become executable in the real world.

The accuracy of current perceptual information can be enhanced based on past memories and associated time and space values. This is because the Perceptual-Multi-Dimensional Memory Frame is simply a memory of object relationships; the Behavior-Multi-Dimensional Memory Frame remembers object relationships associated with the time of the entity's

actions in the procedures of physical behavior; the Relation-Multi-Dimensional Memory Frame stores object relationships associated with procedures and temporal concepts recognized by System 2. There is an overlap between Behavior- and Relation-Multi-Dimensional Memory Frame, and through this overlap, the information in Perceptual-Multi-Dimensional Memory Frame is enhanced.

2) *Basic Senses*: To generate appropriate actions synchronized with the environment in ever-changing situations, it is necessary to bind spatiotemporal information to Multi-Dimensional Memory Frame and render Motor-Multi-Dimensional Memory Frame executable. As a foundational concept to solve this binding problem, the concept of basic senses was proposed, which was identified as the actual manifestation of P-resonance. [10]. P-Resonance occurs between the external stimulus and the Perceptual-Multi-Dimensional Memory Frame by rhythmic and spatial senses. Subsequently, cognitive objects are generated by utilizing the number sense.

a) *Rhythmic Sense*: The changes brought about by actual human action are micro changes on cyclic activity. These minute changes alter the relative situation between the actors themselves and the environment that involves others. However, from the three-dimensional understanding of human perception, these changes are perceived as continuous changes along the time axis. Conversely, if we look at the organs active in the human body, they have evolved and developed under circadian rhythms. Consequently, periodically active organs such as the heart have been formed to provide unique rhythms.

Turning to the environmental side, changes with various reproducible rhythms occur under the cyclic activity of the earth. Thus, to adapt to changes in the environment, a “rhythmic sense of basic perception” should be formed in the connection circuit of the circulatory network formed by various procedural memories. This is called rhythmic sense that enables flexible binding of memory and perceptual information on the time axis in P-Resonance, with hearing as the core and perception in general.

Here, through P-resonance based on rhythmic sense, the region that matches the temporal cyclic pattern of the memory held within the Perceptual-Multi-Dimensional Memory Frame is activated.

b) *Spatial Sense*: Bodily activity includes movement that involves changing the position of one’s own body part in the three-dimensional space. Recognition of the current situation of the three-dimensional space is necessary for constructing executable bodily activities from the information stored in the Motor-Multi-Dimensional Memory Frame, that is free from absolute positions. The unique dimensions associated with movement are distance and time, which are required to make the move. The time is associated with body’s internal rhythms, which define the scale for measuring distance. Thus, the information concerning distance between objects in the external environment is conceived through the rhythm-based scale. This is the second basic sense, the “spatial sense” defined through the rhythmic sense.

Furthermore, the scale can change overtime because human behavior changes its orientation and range of circulation as it grows. Consequently, the cyclic trajectory thus formed evolves into a complicated web reflecting the range of variations of movement; humans expand their activity bandwidth. Inevitably, “spatial sense of basic perception”, which serves as the basis for spatial cognition, should be formed in the connection circuit of the circulatory network formed by various procedural memories, which works in P-Resonance.

Here, procedural memory is formed by overlapping Perceptual-Multi-Dimensional Memory Frame with Behavior- and Relation-Multi-Dimensional Memory Frame, so P-resonance causes activation in that region.

c) *Number Sense*: For humans to select appropriate actions in a timely manner in an ever-changing environment, information related to quantitative comparisons such as larger or smaller for size, more or fewer for the number of objects, farther or closer for distance, and longer or shorter for duration is indispensable. When this information is combined with the reward response that reflects the appropriateness of the choice, a basic sense of quantitative discrimination is formed. This is the third basic sense, the “number sense” [17].

Perceptual information activates Perceptual-, Relation-, and Behavior-Multi-Dimensional Memory Frame through P-resonance mediated by rhythmic and spatial senses. This process takes place in the midst of synchronization between the environment and human activity, which is weak synchronization, synchronization within the width of the time or activity bandwidth [18]. Number sense integrates them into cognitive objects that can be consciously manipulated.

The cognitive objects generated by the number sense are relational networks composed of the following items.

- 1) The Perceptual-Multi-Dimensional Memory Frame activated by the rhythmic sense,
- 2) Existing in its overlapping domains, and associated with Relation-, Behavior-Multi-Dimensional Memory Frame, and Perceptual-Multi-Dimensional Memory Frame activated by spatial sense,
- 3) Symbols existing in the overlapping region of 1 and 2 within the Word-Multi-Dimensional Memory Frame.

The number sense activates these regions, making the Multi-Dimensional Memory Frame available for cognitive processes via C-resonance.

E. C-Resonance

This section explains the mechanism by which C-resonance arises between cognitive processes and Word-, Relation-, Behavior-, and Motor-Multi-Dimensional Memory Frame, based on our previous work [8].

1) *Characteristics of the C-Resonance*: C-resonance occurs between cognitive objects represented as activation patterns in Perceptual-, Behavior-, Relation-, and Word-Multi-Dimensional Memory Frame and System 1 and System 2, driving the cognitive process. As a result, activation propagates to Motor-Multi-Dimensional Memory Frame, mapping to the brain’s representation of motor movement. This is converted

into actionable information through spatiotemporal interpolation during the motor process, and movement occurs via motor nerves. The body plan (skeleton) underlies the entire movement as an interpolation program base, functioning as a default value.

The Multi-Dimensional Memory Frame contains three hierarchically structured memes—C-, B-, and A-memes—as shown in the upper right of Figure 1. These memes correspond to each memory in the Multi-Dimensional Memory Frame. Specifically, C-memes corresponds to Word- and Relation-Multi-Dimensional Memory Frame, B-memes corresponds to Behavior-Multi-Dimensional Memory Frame, and A-memes corresponds to Motor-Multi-Dimensional Memory Frame. C- and B-memes are linked to each other by sharing the Perceptual-Multi-Dimensional Memory Frame. B- and A-memes are related through the overlap between the existence of the former in Behavior-Multi-Dimensional Memory Frame and the existence of the latter in Motor-Multi-Dimensional Memory Frame. These associations form a structure that binds them together. This ensures the reality derived from perceptual information linked to the real world.

2) *Implementation of C-Resonance by GOMS*: C-resonance operates under time constraints where the PCM process must synchronize with environmental changes while selecting and executing actions, linking the Multi-Dimensional Memory Frame with cognitive processes. In our previous work [8] proposing C-resonance, we introduced GOMS, proposed by Card, Moran, and Newell [1], as a mechanism to directly link the Word-, Relation-, Behavior-, and Motor-Multi-Dimensional Memory Frame without going through Perceptual-Multi-Dimensional Memory Frame.

Using memes in an interpreter-like manner via Perceptual-Multi-Dimensional Memory Frame ensures corporeality, but it requires monitoring the situation step by step while advancing the PCM process, making it inefficient. This situation corresponds to MHP/RT running in Mode 2 under System 2 control. Meanwhile, the use of Multi-Dimensional Memory Frame in the form compiled by GOMS corresponds to situations where MHP/RT is running in Mode 1 under System 1 control. This allows us to consider GOMS as implementing C-resonance.

3) *Binding Memes via GOMS*: In GOMS, behavioral goals form a robust hierarchical structure. The goal structure mediates the organization of behavior. Achieving a goal, G , requires achieving the subgoals underneath it, G 's. This structure does not hold the time as its primary parameter. The order between G 's is important. The time elapsed for executing G ' is associated with the operators located at the bottom layer, which connect to the motor process of PCM that implements the contents of Motor-Multi-Dimensional Memory Frame—i.e., the operators—in the real world.

One may reasonably assume that GOMS is used to structure A-, B-, and C-memes which do not contain absolute spatiotemporal information as a method of realizing behavior generation without breaking down, while keeping in sync with the real world where the situation changes from moment to moment. GOMS should correspond to the phenomenon of A-, B-, and

C-memes binding without the Perceptual-Multi-Dimensional Memory Frame when encountering certain situations, indicating the entity of the phenomenon of C-resonance. This may correspond to a shortcut that may be formed within the Multi-Dimensional Memory Frame. Here, GOMS is positioned as a shortcut that guarantees symbol grounding without going through perception.

4) *Symbol Grounding via GOMS*: The embodiment of one's conceptualized ideas in the real world as envisioned—symbol grounding—is crucial for facilitating smooth interaction.

Figure 1 top right shows GOMS substituting the mapping from the Word- to Motor-Multi-Dimensional Memory Frame, where symbol grounding is guaranteed. The self propagates activations through Perceptual-Multi-Dimensional Memory Frame to Relation- and Behavior-Multi-Dimensional Memory Frame, ultimately activating Motor-Multi-Dimensional Memory Frame, for the concept $\hat{C} \in$ C-memes that the self intends to realize in the real world. The goal expressed here is a kind of concept. Then, through the motor process, the object $\hat{O}(\hat{C})$ embodies itself in the real world. What is embodied here is the state in which the goal has been achieved, existing within the environment as a perceivable object. This object is perceived by the self, and the Perceptual-Multi-Dimensional Memory Frame is activated via P-resonance. The activation propagates to Word-Multi-Dimensional Memory Frame, and when the object is recognized as \hat{C} , the relationship, $\hat{C} \equiv \hat{O}(\hat{C})$, holds true, meaning the envisioned concept has been successfully realized in the real world. In other words, it means that the envisioned goal has been realized in the environment. The establishment of this relationship also signifies the resolution of the symbol grounding problem [19], which has been a challenge in the field of artificial intelligence [7].

III. RESONANT ARTIFICIAL OBJECT DESIGN

Figure 3 illustrates that the following occurs for the concept $\hat{C} \in$ the Word-Multi-Dimensional Memory Frame—C-memes—corresponding to the state one intends to achieve. First, the cognitive process resonates with GOMS, which includes the corresponding goal, activating System 2 and System 1, thereby generating the object $\hat{O}(\hat{C})$ in the environment. Subsequently, this object captures the basic senses, inducing P-resonance and forming cognitive objects in the Multi-Dimensional Memory Frame. If the cognitive objects are linked to \hat{O} , symbol grounding is guaranteed. Below are key design considerations to ensure sufficient C-resonance and P-resonance in artificial objects.

A. Key Considerations for Promoting C-Resonance

Since the true nature of C-resonance is GOMS, it is necessary to represent situations where virtual environments are utilized using GOMS. This is achieved by constructing a GOMS representation of human behavioral ecology within the target virtual environment. GOMS organizes C-, B-, and A-memes into a hierarchical structure as shown in Figure 3. Therefore, the procedure for constructing GOMS begins with extracting

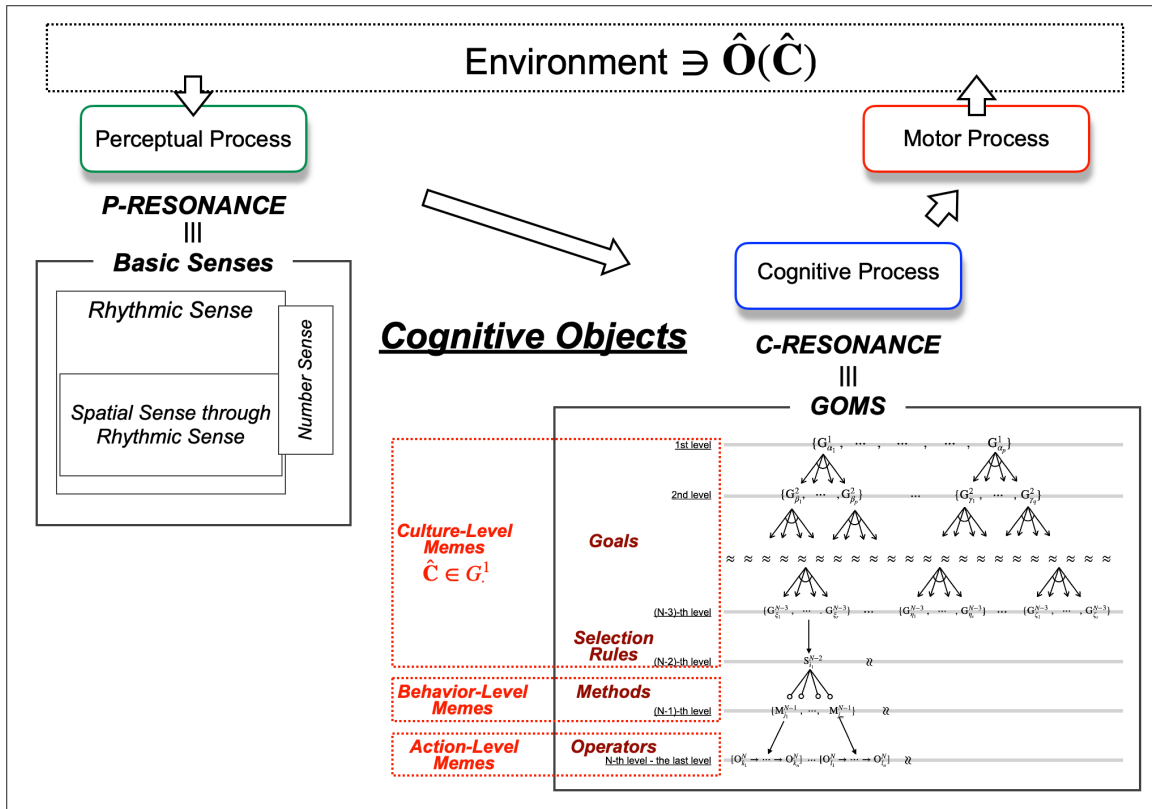


Figure 3. Activation of environment-synchronized PCM processes and environment-asynchronous Multi-Dimensional Memory Frame—P-resonance via basic senses, C-resonance via GOMS, and the cognitive objects connecting them.

memes that are inherited across generations within the target behavioral ecology. Here, a method known as Cognitive Chrono-Ethnography (CCE) [20] proves effective. This involves conducting behavioral simulations using MHP/RT to identify parameters characterizing behavioral patterns, observing human actions exhibiting distinctive combinations of these parameter values, and thereby providing detailed descriptions of behavioral ecology. This method has been implemented in several case studies [21][22].

A-memes are observable. Mutually distinguishable A-memes define the operators “O”. B-memes are pattern sequences associated with methods “M”, which are patterns reproducibly generated when conditions are met by sequences of A-memes. The conditions define selection rules “S”, expressed as “IF condition THEN pattern sequence name” consciously processed by System 2. Upon this, goal structures unfold.

At the top, goals related to happiness and satisfaction—such as winning competitions and gaining intellectual fulfillment—are placed, with the hierarchy unfolding downward. A goal structure is expressed by linking goals represented as concepts in the form of goals and subgoals. Note that a higher-level goal is achieved by accomplishing all the lower-level subgoals connected to it.

GOMS does not consciously evaluate the results of actions because higher-level goals are set by System 2, and System 1 executes actions via feedforward control—execution in Mode 1.

To promote C-resonance, it is necessary to accurately grasp in advance the goal structure that humans are attempting to achieve, identify the goal being pursued from the observed sequence of actions, and then appropriately infer the next goal that will likely be set.

B. Key Considerations for Promoting P-Resonance

The rhythmic sense, spatial sense, and number sense are applied to generate cognitive objects from the resonated portion of the Perceptual-Multi-Dimensional Memory Frame for further cognitive processing carried out by System 1 and System 2. The accuracy of adaptive behavior can be increased by repeating the behavior while developing the rhythmic, spatial, and number senses in the reuse of the Perceptual-Multi-Dimensional Memory Frame. Objects in the environment \hat{O} serve as the starting point for P-resonance, determining the activation patterns of the Perceptual-Multi-Dimensional Memory Frame through rhythmic, spatial, and number senses. The content of objects that individuals can interact with within their limited time influences the development of basic senses. Therefore, it is necessary to design objects \hat{O} in the environment to elicit responses through basic senses.

1) *Conditions that Trigger Rhythm Sense:* Humans form the rhythmic sense as fundamental perception in the connection circuits of a cyclic network formed by various procedural memories in order to adapt to environmental changes. A

perceived rhythm is a subjective grouping of events occurring successively along the time axis generated through the rhythmic sense, and its nature is primarily determined by the length of the intervals between events. When events recur periodically according to a certain pattern, the temporal structure of that pattern is recognized as rhythm. The most fundamental element of rhythm is the existence of regularity or repetitive patterns in the intervals between events. Through this regular interval, humans can predict what will happen next and perceive it as a coherent whole. The regularity and periodicity of the intervals between events play a central role in perceiving and defining rhythm. Therefore, the events arising from interaction generating patterns along the time axis become a necessary condition for rhythm perception to be triggered.

2) *Conditions that Trigger Spatial Sense:* The determination of rhythm signifies that the unit time \hat{T} characterizing events occurring in the environment is established. Event $E(T)$ occurring at time T becomes comparable to an event occurring at time $T + \hat{T}$, and the difference $D(E(T + \hat{T}), E(T))$ between the events becomes discernible. When events occurring in three-dimensional space are the subject, spatial cognition—the cognitive representation of space—is generated based on perceived differences. This spatial cognition is referenced when recognizing and predicting changes in the relative position between the environment and oneself over time.

Objects in the environment can evolve over time within three-dimensional space according to their own programmed dynamics and generate events. However, it is necessary to consider that the receiving human side quantizes events according to a rhythm characterized by the unit time \hat{T} .

3) *Conditions that Trigger Number Sense:* By utilizing spatial sense, the temporal development of the relative relationship between the environment and the self is generated as activation patterns of Perceptual-, Behavior-, and Relation-Multi-Dimensional Memory Frame. To advance the situation toward the immediate behavioral goal, it is necessary to consolidate its activation patterns into a symbol and evaluate them against the behavioral goal defined by System 2. Number sense is used to select information concerning quantitative comparisons of about three items, which is necessary for choosing appropriate actions at the right time in an environment that changes moment by moment. The evaluation can be conducted in a multifaceted manner, but comparisons involving three or fewer items are preferable. The parameters characterizing programmed changes in the external environment must be linked to the evaluation of whether the goal has been achieved.

IV. CONCLUSION AND FUTURE WORK

It is desirable for people to strive toward achieving their goals, interact with their surrounding environment, evaluate the results, and gain a sense of accomplishment. When considering a situation where the environment returns pre-programmed responses to human actions, the environment can be regarded as virtual in the sense that it is not the actual environment. As a result of one's own actions, some objects are returned from the virtual environment. If one can identify the content of their

own actions within these objects and use it to set the content for the next action, the interaction between the human and the virtual environment can be considered to be progressing smoothly.

Interaction is executed through a PCM process that synchronizes with the ever-changing environment. The execution content is determined by the activation patterns of the Multi-Dimensional Memory Frame, which operate independently of environmental changes. To maintain smooth interaction, it is important to control the content of P-resonance generated by basic senses and the content of C-resonance generated by GOMS. This paper has outlined the matters that should be considered in these controls.

Guidelines are generally expressed in abstract terms, so to apply them to real-world artificial environment design scenarios, implementation methods for the guidelines must accompany them. This study showed that designs aiming to induce P-resonance and C-resonance, and designs enabling the functioning of basic senses and GOMS, are effective as practical implementations of the guidelines [13] proposed in relation to the MHP/RT operating modes (specifically Mode 1) shown in Figure 2. Whether the basic senses function can be examined by providing representations for the perceptual information provided. GOMS can be examined by representing the knowledge required for action selection. Each of these is fundamentally based on static analysis.

In P-resonance, it was pointed out that it is important to consider that the activation of Perceptual-, Relation-, and Behavior-Multi-Dimensional Memory Frame originating from perceptual information occurs in situations where the environment and human activity are weakly synchronized. When a system and a human resonate, the two systems intertwine and synchronize at a specific point in time. Meanwhile, weak synchronization [18][23] is a concept of synchronization applied to human activities occurring within a temporal or activity bandwidth, extending the aforementioned concept of synchronization. Specifically, actions executed at time T involve *System 2* and *System 1* executed beforehand by MHP/RT, and *System 1* and *System 2* executed afterward. These four processing procedures synchronize with the system event at time T . This synchronization method is called weak synchronization. We also propose design guidelines related to weak synchronization [13].

Just as practice became possible by focusing on basic senses and GOMS—concepts external to MHP/RT—the next challenge is to make artifact design guidelines, derived from weak synchronization (the conditions for P-resonance) associated with MHP/RT's four processing modes, practicable by introducing concepts external to MHP/RT. Here, dynamic analysis focusing on the nature of interactions over time is fundamental. For example, the degree of weak synchronization in situations where latent information about visual objects presented in a VR learning environment is provided auditorily can be examined by measuring visual behavior and pupil responses, and analyzing the relationship between these results and memories of audiovisual experiences [24][25].

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI (Grant Numbers 19K12246 / 20H04290 / 22K12284 / 23K11334) and the National University Management Reform Promotion Project.

REFERENCES

- [1] S. K. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.
- [2] A. Newell, "Physical Symbol Systems", *Cognitive Science*, vol. 4, no. 2, pp. 135–183, 1980. DOI: 10.1207/s15516709cog0402_2[retrieved:February,2026].
- [3] A. Newell, *Unified Theories of Cognition (The William James Lectures, 1987)*. Cambridge, MA: Harvard University Press, 1990.
- [4] J. E. Laird, A. Newell, and P. S. Rosenbloom, "Soar: An architecture for general intelligence", *Artificial Intelligence*, vol. 33, pp. 1–64, 1987.
- [5] M. Kitajima and M. Toyota, "Decision-making and action selection in Two Minds: An analysis based on Model Human Processor with Realtime Constraints (MHP/RT)", *Biologically Inspired Cognitive Architectures*, vol. 5, pp. 82–93, 2013, ISSN: 2212-683X. DOI: 10.1016/j.bica.2013.05.003[retrieved: February,2026].
- [6] M. Kitajima, *Memory and Action Selection in Human-Machine Interaction*. Wiley-ISTE, 2016, ISBN: 9781848219274.
- [7] M. Kitajima, M. Toyota, and K. T. Nakahira, "Why the Symbol Grounding Problem Matters in Virtual Reality: A Meme-Focused Solution Based on the Model Human Processor with Real-Time Constraints", *International Journal On Advances in Intelligent Systems*, vol. 18, no. 3 & 4, pp. 162–172, 2025.
- [8] M. Kitajima, M. Toyota, J. Dinet, and K. T. Nakahira, "Transforming Conscious Goals into Unconscious Actions in Real-world Interactions: Real-world Use of Behavioral Ecological Memes via GOMS", *International Journal On Advances in Intelligent Systems*, vol. 18, no. 3 & 4, pp. 173–186, 2025.
- [9] M. Kitajima and M. Toyota, "Simulating navigation behaviour based on the architecture model Model Human Processor with Real-Time Constraints (MHP/RT)", *Behaviour & Information Technology*, vol. 31, no. 1, pp. 41–58, 2012. DOI: 10.1080/0144929X.2011.602427[retrieved:February,2026].
- [10] M. Kitajima et al., "Basic Senses and Their Implications for Immersive Virtual Reality Design", in *AIVR 2024 : The First International Conference on Artificial Intelligence and Immersive Virtual Reality*, 2024, pp. 31–38.
- [11] D. Kahneman, "A perspective on judgment and choice", *American Psychologist*, vol. 58, no. 9, pp. 697–720, 2003.
- [12] D. Kahneman, *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux, 2011.
- [13] M. Kitajima, M. Toyota, and J. Dinet, "Guidelines for Designing Interactions Between Autonomous Artificial Systems and Human Beings to Achieve Sustainable Development Goals", *International Journal on Advances in Intelligent Systems*, vol. 15, pp. 188–200, 2022.
- [14] M. Toyota, M. Kitajima, and H. Shimada, "Structured Meme Theory: How is informational inheritance maintained?", in *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, 2008, p. 2288.
- [15] D. C. Dennett, *From Bacteria to Bach and Back: The Evolution of Minds*. W W Norton & Co Inc, Feb. 2018.
- [16] M. Kitajima, M. Toyota, and J. Dinet, "How Resonance Works for Development and Propagation of Memes", *International Journal On Advances in Systems and Measurements*, vol. 14, pp. 148–161, 2021.
- [17] S. Dehaene, *The Number Sense: How the Mind Creates Mathematics, Revised and Updated Edition (English Edition)*. Oxford University Press, Apr. 29, 2011, p. 604.
- [18] J. Dinet and M. Kitajima, "Immersive interfaces for engagement and learning: Cognitive implications", in *Proceedings of the 2015 Virtual Reality International Conference*, ser. VRIC '18, Laval, France: ACM, 2018, 18/04:1–18/04:8. DOI: 10.1145/3234253.3234301[retrieved:February,2026].
- [19] S. Harnad, "The symbol grounding problem", *Physica D: Nonlinear Phenomena*, vol. 42, no. 1, pp. 335–346, 1990, ISSN: 0167-2789. DOI: 10.1016/0167-2789(90)90087-6[retrieved: February,2026].
- [20] M. Kitajima, "Cognitive Chrono-Ethnography (CCE): A Behavioral Study Methodology Underpinned by the Cognitive Architecture, MHP/RT", in *Proceedings of the 41st Annual Conference of the Cognitive Science Society*, Cognitive Science Society, 2019, pp. 55–56.
- [21] M. Kitajima, M. Toyota, and J. Dinet, "Art and Brain with Kazuo Takiguchi - Revealing the Meme Structure from the Process of Creating Traditional Crafts -", in *COGNITIVE 2023 : The Fifteenth International Conference on Advanced Cognitive Technologies and Applications*, 2023, pp. 1–10.
- [22] K. T. Nakahira, M. Kitajima, and M. Toyota, "Understanding Practice Stages for a Proficient Piano Player to Complete a Piece: Focusing on the Interplay Between Conscious and Unconscious Processes", *International Journal on Advances in Life Sciences*, vol. 16, pp. 164–177, 2024.
- [23] M. Kitajima, J. Dinet, and M. Toyota, "Multimodal Interactions Viewed as Dual Process on Multi-Dimensional Memory Frames under Weak Synchronization", in *COGNITIVE 2019 : The Eleventh International Conference on Advanced Cognitive Technologies and Applications*, 2019, pp. 44–51.
- [24] Y. Kurihara, M. Shino, K. Nakahira, and M. Kitajima, "Visual behavior based on information foraging theory toward designing of auditory information", in *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 1: HUCAPP*, INSTICC, SciTePress, 2024, pp. 530–537. DOI: 10.5220/0012474600003660[retrieved:February,2026].
- [25] R. Hirabayashi., M. Shino., K. N. T., and M. Kitajima., "How auditory information presentation timings affect memory when watching omnidirectional movie with audio guide", in *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 2: HUCAPP*, INSTICC, SciTePress, 2020, pp. 162–169. DOI: 10.5220/0008966201620169[retrieved: February,2026].

C-LISTEN: Cognitive Load Impacts of Sensory-Triggered Environmental Navigation in Virtual Reality

Md. Monowar Hossain 

Department of Computer Science
Kennesaw State University, United States
e-mail: murad0904045@gmail.com
mhossa36@students.kennesaw.edu

M. Rasel Mahmud

Assistant Professor of Computer Science
Kennesaw State University, United States
e-mail: m.raselmahmud1@gmail.com
mmahmud2@kennesaw.edu

Abstract—Nowadays, Virtual Reality (VR) systems are more helpful for making decisions, training, and rehabilitation. Multimodal settings in these systems impose a significant balance and cognitive difficulties. Additionally, decreased performance, higher cognitive load, user overwhelm, and limited accessibility of VR technology can be caused by excessive auditory, visual, and sensory stimulation. In our pilot study, we examine the strategic design of directional auditory cues to improve cognitive load and enhance user experience in virtual environments. This research also enhances fundamental knowledge that auditory feedback reduces cognitive load in VR, with statistical analysis confirming this improvement ($p = .0011$). This study investigates how postural balance and pupil diameter correlate with cognitive load while the users navigate. Moreover, this study provides realistic overview design guidelines for accessibility, which will make VR experiences safer and more useful for everyone, including people with cognitive impairments.

Keywords-Cognitive Load; Extended Reality; Virtual Reality; Auditory Feedback.

I. INTRODUCTION

Virtual Reality (VR) has the ability to create interactive and immersive environments that are used in different sectors, such as healthcare, training, education, and entertainment. Cognitive load refers to the amount of mental effort for processing information while accomplishing a task, which is strongly influenced by the effectiveness of VR experiences. One of the major reasons for ineffective design is that it leads to a degradation of performance, discomfort, and overload of users' cognitive resources.

The cognitive load is related to the participants' pupil diameter and body balance. Kahneman et al. [1] found that when a participant processed information in short-term memory, the pupil diameter was larger than in a normal situation. Pupil size increased when performing a difficult task, which reflects the amount of mental effort or cognitive load on memory. The impact continues even when visual factors remain controlled, which indicates that pupil dilation is associated with cognitive processing rather than only adjustments in eye focus. In general, the size of the pupil is a good predictor of how much memory and mental work a person is utilizing at any given time.

In Small et al.'s work [2], participant body balance is degraded during the performance of a listening or spelling task involving short and long words, as well as the walking dual task condition. The researchers established that the participant prioritized the thinking task over balance, and a higher cognitive load led to a decline in balance control during

walking. In [3], without a feedback system in immersive VR, the standing balance generally declined, and adding assistive feedback improved the balance. They noticed that individual assistive feedback (auditory, vibrotactile, or visual) can improve balance without noticeably increasing cognitive load than the multimodal condition.

We investigated 15 types of audio, categorized into four types (classical, natural, electronic, and unnatural), which helped to reduce cognitive load. In the classical music conditions included four common music from Western classical music: Adagio in G minor (Albinoni), Allegro con spirito from Sonata K.448 (Mozart), Gymnopédie No. 1 (Satie), and Für Elise (Beethoven). These musics were chosen because they are common for listening conditions in cognitive research [4]–[9]. The natural soundscape had a common combination of sounds, including fountains and birds tweeting, ocean waves, and birdsong. Researchers often use them to examine how they can help people feel less stressed and less mentally burdened [10][11]. High-energy tracks, such as as Balearic Pumping, Natural (Imagine Dragons), Rum n' Bass (Boom Kitty), and POP/STARS (K/DA) are known as electronic music. Strong rhythmic drive, amplified bass, and fast, repetitive beats are features included in this type of music that affect arousal and attentional activation. During the task, these were used to investigate how they may influence focus and cognitive load [12]. There are commonly unnatural sounds like aviation noise, motorcycle noise, white noise, and construction noise used that represent the human-made acoustic environment in research to examine how they can impact stress and cognitive load [3][11][13].

The key contributions of our research include: First, this study investigates whether assistive auditory cues can help users manage cognitive load in virtual reality (VR) environments. The results show that well-designed auditory guidance can help users control their mental effort during VR navigation tasks. Second, the study analyzes how different categories of audio cues affect cognitive load. This highlights the importance of selecting appropriate auditory feedback when designing VR systems. Third, this work examines the relationship between postural balance and cognitive load. The findings suggest that higher cognitive load can affect users' physical stability during VR interaction. Finally, the study identifies a correlation between pupil diameter and cognitive load, indicating that pupil dilation can be used as a physiological measure to monitor

users' mental effort in immersive environments. Overall, these findings provide useful insights into how auditory feedback influences both cognitive and physical responses in VR, which can help in designing safer and more efficient VR systems.

The rest of the paper is explained as follows: In Section II, a background study is discussed, which contains the previous research summary of the assistive auditory feedback system done by other researchers. In Section III, we describe our proposed methods. Lastly, Sections IV, V, and VI detail the statistical analysis & experimental results with discussion and conclusion, respectively.

II. BACKGROUND STUDY

This section reviews existing literature on assistive auditory feedback systems and their role in influencing cognitive load, user performance, and navigation in virtual environments.

A. Assistive Auditory Feedback Technology

While improving the efficiency of task completion, controlling cognitive load is one of the vital goals of assistive auditory feedback technology, which provides information through audio. Well-designed auditory feedback cues, such as earcons, auditory icons, speech prompts, and sonification, can enhance user performance and support usability, which is related to reducing unnecessary cognitive effort [14].

Multimodal feedback can reduce task difficulty in VR tasks [15][16]. To find out how VR-based assistive auditory, visual, vibrotactile, or multimodal feedback can help people with balance and gait problems in [17]. Furthermore, Mahmud et al. [3] found that cognitive load is higher ($p = .04$) when multimodal feedback is used than when individual feedback is used. Ricci et al. also found that a single modality, like auditory or haptic cues, is more effective than multimodal signals due to helping to easily distribute attention in complex situations for navigation. Auditory feedback can enhance learning and memory utilization by reinforcing task-relevant cues [18] and improve postural balance [19]. Real-world studies demonstrate that auditory assistive cues can improve navigation, attention management, and error correction [19][20].

Although complex sound can affect the cognitive load for the navigation system, for those who have a problem with visualization, the auditory feedback helps to navigate through detecting obstacles and plotting a path [21]. For a safety-critical environment like automated driving, the "speech" and "spearcon" cues performed better than earcons, as they helped to understand the critical situation faster and solve the situation using less mental stress [22]. Bruckman et. al. [23] analysed that helps cognitive support in IDEs using the auditory feedback. This auditory feedback provides information on code syntax and error details in programming tasks for solving problems.

1) *Classical Music as an Auditory Feedback*: Mozart's Sonata is the most popular piece used for research in music, cognitive and brain activity. The authors examined the past 30 years of studies on humans and animals that revealed how the person's excitement, preferences, and listening context affected their memory, attention, and brain activity [4]. Mozart's Sonata

and Adagio in G Minor music were used in the study [5], and the researcher examined how these influenced the working memory in different age groups. They also found that there is no overall effect on complex tasks. In Xing et al.'s work [6], for improving learning and task performance, the music of "Mozart's Sonata" is a better solution than other versions that have increased performance and provide a more positive effect on brain activity.

In Jauvsovec et al.'s study [7], while solving spatial rotation tasks and math tasks, the brain activity is affected by listening to Mozart's music. This music degraded the performance of the math task, but it provided a better performance in the spatial task. The researchers analyzed that the music can boost the activity in the brain area that assists in combining the information into meaningful. Shen et al. [8] investigated how the soft background music (Gymnopedie No.1) without lyrics affects participants' attention and showed that this music performs better than silence. They also suggested that soft background music can improve participant attention and quick learning. Mohammadian et al. [9] explored how listening to "Für Elise" music in noisy simulated open-plan offices contributed to improving the mental workload more than using irrelevant speech. They used n-back tests for assessing performance accuracy and provided suggestions that the background music can act as a counteraction to noise and provide more support for cognitive performance. We chose auditory feedback over vibrotactile [24]–[26] or visual [27][28] feedback because participants had to wear vibrating devices for vibrotactile feedback and had to look at visual signs for visual feedback, which may increase their cognitive load.

2) *Natural Music as an Auditory Feedback*: The author [10] examined that a combination of sounds from a fountain and tweeting birds reduces the stress level than noisy sounds. Moreover, nature sounds can restore skin conductance levels more quickly after completing the stressful task. They also observed that for participants' relaxation, natural sounds play a vital role. After stressing and during rest, they [11] analyzed how different natural sounds, such as "ocean waves and birds' songs" affect the participants' mental workload. They also found that heart rate decreases after completing a stress task and that skin conductance increases when participants hear natural sounds during rest, helping reduce cognitive load.

3) *Electronic Music as an Auditory Feedback*: The authors investigated [12] how musical training affected users' cognitive load and task performance during VR Beat Saber gameplay. In the musical training, they used the common music for this game: "Balearic Pumping, Natural, Rum n' Bass, and POP/STARS". In addition, they found that the harder level of the task required more mental effort, but musical training helped to improve cognitive load.

4) *Unnatural Music as an Auditory Feedback*: In Medvedev et al. study [11], aviation, motorcycle, and construction sounds are deliberated as unnatural sounds that negatively affect the participants' perceptual and physiological strain. Participants generally considered these sounds as calming, restorative, and these sounds were compared with natural sounds, birds' songs,

and ocean waves. The findings of this study were that these sounds have an impact on greater physiological strain (higher skin conductance) and a slow rate of stress recovery. Moreover, they suggested that it may elevate the cognitive load.

The researchers [13] wanted to know whether the addition of white noise could affect children's thinking and attention with different levels of ability. They suggested, on the basis of the Moderate Brain Arousal model, every person has a different amount of natural neural noise, and improving the children's brain activity may improve when adding external white noise. Mahmud et al. [3] also included white noise in their study, which was used in multimodal feedback methods for enhanced participant stability and their comfort.

III. METHODS

This section outlines the experimental design, including system setup, participant details, auditory conditions, and procedures used to investigate the relationship between auditory feedback and cognitive load in VR navigation.

A. System Description

1) *Balance Measurement*: In this study, we employed the BTrackS Balance Plate, which operates at a 25 Hz sampling frequency. This device was used to collect different types of balance, such as sway and stability, which helped to analyze the distribution of pressure applied to the plate. Ensuring high-resolution tracking of participants' balance metrics data helped to identify and measure cognitive load for each scenario.

2) *Safety Equipment*: We utilized a harness system to ensure participants' safety. When a participant conducts the VR task, there is a chance of falling, so to mitigate this risk, we used this system.

3) *Computers, VR Equipment, and Software*: We had to set up the HTC VIVE Focus Vision, which is characterized by a 100-degree field of view, to ensure immersive and smooth visual experiences and collect the participant's Eye tracking data. The Virtual Environment (VE) rendering and data collection were managed by using a high-performance computer that was configured with an Apple M3 chip, 8-core CPU, 10-core GPU, and RAM 8GB. Using the NI LabVIEW software (version 2020), which helps to collect the BTrackS Balance Plate data on our computer.

4) *Environment*: We performed our task in the eXtended Reality and Intelligence (XRei) Lab at Kennesaw State University's College of Computing and Software Engineering, which was more than 1000 square feet in size.

5) *Participants*: Firstly, we circulated information about participating in our pilot study. Then, we collected all demographic information from those who were interested. As we ran the pilot study, we recruited only six (06) participants with adults aged 18 and older from diverse backgrounds. The detailed information of the participants has been shown in the Table I.

TABLE I. DESCRIPTIVE STATISTICS FOR PARTICIPANTS

Participants		Age (years)		Height (cm)		Weight (lb)	
Male	Female	Mean	SD	Mean	SD	Mean	SD
4	2	27	2.61	169.67	9.37	183	27.71

B. Experimental Conditions

1) *Auditory Feedback*: In this study, we utilized 15 directional auditory cues to provide auditory feedback that helps to investigate the cognitive load level while enhancing the balance and pupil diameter within VR-based navigation tasks. We were categorized into four types of audio (classical, natural, electronic, and unnatural) among 15 audio files.

2) *No Feedback in VR*: To serve as a control condition, the participant played the task in VE without the involvement of any additional audio feedback. To maintain consistency in the environment, participants wore all the necessary equipment and adhered to other experimental conditions. This result was used as a baseline condition for comparison with all other conditions.

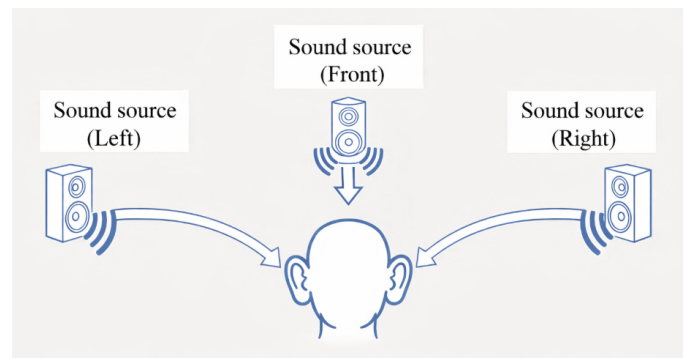


Figure 1. Auditory feedback system

3) *Auditory Feedback Setup*: To ensure precise direction-based three-dimensional sound localization, the sound was provided to the high-fidelity headphones. When a user explored the path and made the decision in which path he selected, our system provided a directional auditory cue. Auditory feedback in the left and right ears indicated to turn left and right, respectively, while the auditory feedback originating from the front indicated to move forward, like as Figure 1. This feedback helped to guide which path was best suitable for the destination goal.

C. Hypotheses

Based on the literature analysis in Sections I and II and recent studies that discuss how extra auditory feedback systems can cause cognitive overload, we suggest the following hypothesis for this research:

- H1: Our hypothesis is that without assistive feedback in VR, the cognitive load in navigation tasks can increase. When a participant performs a task without any feedback, it will increase the cognitive load [29].
- H2: We hypothesize that the participant's postural balance is related to the brain activity when he performs a task. Small et

- al. [2] analyzed that when a participant performs a cognitive task while walking, their ability to maintain dynamic balance decreases under a higher cognitive load.
- H3: Our hypothesis is that as the pupil diameter increases, it also indicates that the participant's cognitive load is high. While a participant performs a task and makes a decision by taking more time, the pupil diameter is increased. Pupil diameter is more related to measuring or indicating cognitive load during a task [1].
 - H4: Participants exposed to classical and natural category audio will experience significantly lower cognitive load than when exposed to electronic and unnatural category audio.
 - H5: There will be a significant difference in cognitive load across the different audio categories (classical, natural, electronic, and unnatural).

D. Metrics

1) *Mental Load Assessment*: A subjective rating system was used in our system to collect the mental load information provided by the participants. After each task completion, each participant was asked to fill out the NASA TLX form on a scale from 0 to 10. From this self-reported evaluation form, we could gauge the cognitive load with each condition. Using this numerical rating system allowed us to measure cognitive load and compare it with other performance metrics and conditions.

2) *Center of Pressure (CoP) Velocity*: The term CoP velocity refers to the information about the dynamic shifts of the body's center of pressure during the participant's involvement in various tasks or conditions. For each condition, this metric was captured systematically and utilized to monitor and measure the changes in balance. After that, it helped to correlate with the participant's cognitive load.

3) *Simulator Sickness Questionnaire (SSQ)*: We utilized the SSQ in our system, which helps to analyze the prevalence of cybersickness and its potential impact on cognitive load and postural stability [30]. To assess the physiological discomfort and symptoms, this SSQ was designed with 16 questions that were collected after the session. Our system analyzed this information to identify the participants' cybersickness and also investigated how these symptoms are related to the cognitive load and balance performance.

E. Study Procedure

1) *Experimental Setup*: The study was approved by the Institutional Review Board (IRB). The participant wore the safety harness and headset to ensure stability and safety throughout the sessions, and placed the balance board in the center position of the safety harness. Participants' balance was collected by using a balance board, and head movement and eye tracking data were stored using an HMD headset during all conditions. Balance signal, head movement, and eye tracking were all recorded as numerical measurements, which were processed for analysis.

2) *Pre-Session Questionnaires*: The Activities-specific Balance Confidence (ABC) questionnaire [31] and the SSQ [30] were answered by the participants after finishing the whole

session. Moreover, in each trial user rated the NASA TLX form on a scale from 0 to 10 for measuring the cognitive load.

3) *Experimental Tasks*: Participants performed a maze navigation task that was conducted in virtual environments. To control for potential order effects, the task order and navigation-oriented auditory feedback were randomized and counterbalanced.

4) *Baseline Measurements*: Participant conducted the maze navigation task without any auditory feedback and captured all information, such as eye tracking, head movement, balance, and the user's mental assessment data, as baseline measurements. This trial lasts about one minute. Finally, these baseline measurements helped to compare with all other conditions and decide whether to correlate with each other.

5) *Tasks in Virtual Reality*: The VR task is similar to the baseline activity, including the additional navigation-oriented auditory feedback conditions. These feedback conditions were randomized and counterbalanced, lasting around one minute.

6) *Maze Navigation in VR*: The VR Maze Game was chosen as the experimental setting because navigating a maze requires various cognitive functions, such as spatial reasoning, working memory, attention control, and decision-making, all of which are important to cognitive load. To navigate a maze, participants must constantly encode and remember spatial information, plan their paths, track their mistakes, and adjust to varying audio cues [32]. All of these things naturally cause changes in mental workload that can be measured.

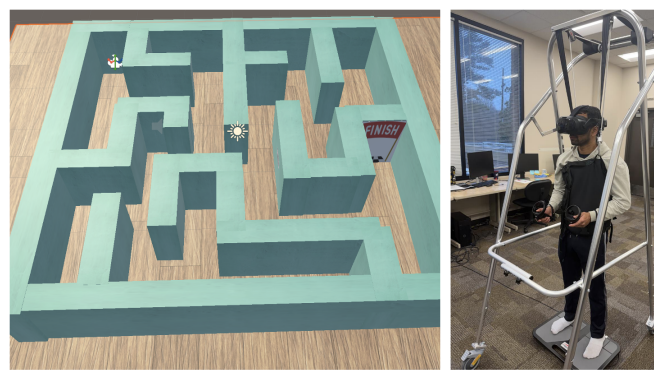


Figure 2. a) Top-down view of the virtual maze environment. b) Participants navigating the virtual maze while standing on the balance board and wearing a harness for safety.

The participants were placed in a flat space during the maze task, which is shown in the Figure 2. In this maze navigation task, the participant used the HTC VIVE Focus Vision headset. Using HTC Vive controllers, the participant moved and turned in the maze environment. The left controller was used for moving, and the right controller was used for turning in the path. When a user explored the path and made the decision in which path he selected, our system provided directional auditory cues. Auditory feedback in the left and right ears indicated to turn left and right, respectively, while the auditory feedback originating from the front indicated to move forward. This feedback helped to guide which path was best suitable

for the destination goal. The entire session will be conducted based on navigation-oriented auditory feedback, and it took 1-2 minutes to navigate mazes.

7) *Post-Session Questionnaires*: After each condition, participants completed the NASA-TLX survey to measure cognitive load. Following the completion of all conditions, users were asked to fill out two distinct questionnaires: the SSQ and a demographic questionnaire. These questionnaires played a vital role in collecting all information related to simulator sickness experienced and the relevant background information on the users.

IV. STATISTICAL ANALYSIS & RESULTS

This section presents the statistical methods applied to the collected data and discusses the key findings regarding cognitive load, balance, and pupil diameter across different auditory conditions.

For the statistical analysis, ANOVA and t-test were used in different studies [3][17]. We utilized a mixed model ANOVA for our experimental records. For this study, we designed a 3×5 factorial structure that evaluates the impact of three independent variables (balance, cognitive load, and pupil diameter), each with multiple participants, on the dependent variables. For further analysis of the ANOVA test, we performed pairwise comparisons using two-tailed t-tests. These comparisons were measured for differences between all combinations of study conditions. From these outcomes, we can determine statistical significance, thereby boosting the robustness of our findings.

A. Effect of Auditory Feedback on Cognitive Load

The repeated-measures ANOVA provided a significant difference in the impact of different auditory feedback on cognitive load, $F(4, 20) = 7.02$, $p = .0011$. The result demonstrated that classical and natural audio feedback conditions reduced the cognitive load more than the no-audio feedback condition. Classical Music (Mean = 2.12) and Natural Music (Mean = 2.34) have lower cognitive value than Electronic Music (Mean = 2.89), and the highest is for Unnatural Music (Mean = 3.27). Thus, the result supported our hypothesis H5.

B. Pairwise Comparisons Against No-Audio Baseline

We showed in Table II each sound condition with the no-audio baseline using a t-test. This analysis demonstrated that classical music ($t=-2.47$, $p=.057$, $dz=-1.01$) and natural music ($t=-2.33$, $p=.068$, $dz=-0.95$) presented marginally significant improvements in cognitive load. On the other side, Electronic Music ($t=-1.40$, $p=.221$, $dz=-0.57$) and Unnatural Music ($t=0.16$, $p=.875$, $dz=0.07$) had no meaningful difference compared to the baseline condition. This analysis established that classical and natural audio categories have a greater effect on cognitive load than other categories, thus supporting H4.

C. Relationship Between Cognitive Load and Balance

The correlation analysis in Table II pointed out that cognitive load and postural balance have a reverse relationship. Classical Music ($r=-0.82$, $p=.048$) and Natural Music ($r=-0.98$, $p<.001$)

have statistically significant results, while Electronic Music ($r=-0.47$, $p=.351$) and Unnatural Music ($r=-0.72$, $p=.107$) showed weaker and non-significant trends in Figure 3. Therefore, it confirmed that a higher cognitive load is associated with lower postural balance, supporting H2.

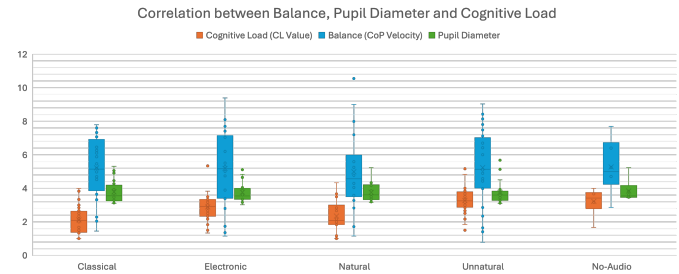


Figure 3. Correlation between Balance, Pupil Diameter and Cognitive Load for Different Types of Auditory Feedback

D. Relationship Between Cognitive Load and Pupil Diameter

The table indicates a positive relation between pupil diameter and cognitive load across all auditory feedback conditions. For the classical category ($r=0.872$, $p=0.0236$), there is a statistically significant correlation, while natural music ($r=0.803$, $p=0.0543$) has a marginal relationship. On the other hand, electronic ($r=0.555$, $p=0.2532$) and unnatural categories ($r=0.766$, $p=.0758$) exhibited non-significant trends. Overall, these results indicated that higher cognitive load tends to be associated with increased pupil diameter in Figure 3, which partially supported H3. Though only the classical category fills the conventional levels of statistical significance.

E. Assistive Auditory Feedback vs No-Auditory Feedback

From the comparison of cognitive load between no-audio feedback and different categories of assistive auditory feedback, the classical and natural categories provide better performance in Figure 3. Since cognitive load was strongly related to balance, these reductions indirectly helped maintain postural stability. So, assistive auditory feedback can make it easier to navigate tasks in a VR environment that is supported by H1.

F. Simulator Sickness Questionnaire

We calculated a two-tailed t-test between pre-session SSQ score and post-session SSQ score of participants. We found the statistically significant difference between pre-session SSQ and post-session SSQ of the participants. We got the result $t(5) = 4.6377$, $p = .005644$, Cohen's $dz = 1.89$ for participants.

V. DISCUSSION

This section interprets the experimental results, highlighting their implications for cognitive load management, balance control, and the design of effective auditory feedback in VR systems.

TABLE II. CORRELATION BETWEEN BALANCE, PUPIL DIAMETER AND COGNITIVE LOAD FOR VARIOUS AUDITORY CONDITIONS

Comparison	t_{CgLd}	p_{CgLd}	d_z^{CgLd}	$t_{Bal.}$	$p_{Bal.}$	$d_z^{Bal.}$	t_{Pupil}	p_{Pupil}	d_z^{Pupil}	$r_{(Bal,CgLd)}$	$Pr_{(Bal,CgLd)}$	$r_{(Pupil,CgLd)}$	$Pr_{(Pupil,CgLd)}$
Classical													
vs No-Audio	-2.47	0.0568	-1.01	-0.11	0.918	-0.04	-0.50	0.638	-0.20	-0.82	0.0480*	0.872	0.0236*
Electronic													
vs No-Audio	-1.40	0.2209	-0.57	-0.17	0.874	-0.07	-1.04	0.347	-0.42	-0.47	0.3507	0.555	0.2532
Natural													
vs No-Audio	-2.33	0.0676	-0.95	-0.46	0.666	-0.19	-0.62	0.564	-0.25	-0.98	0.00048*	0.803	0.0543
Unnatural													
vs No-Audio	0.16	0.875	0.07	-0.06	0.954	-0.02	-2.75	0.0401*	-1.12	-0.72	0.1072	0.766	0.0758

Note. t = paired-sample t statistic; p = probability value; d_z = Cohen's effect size for paired samples; r = Pearson correlation between balance, pupil diameter and cognitive load. * indicates $p < .05$.

A. Effect of Assitive Auditory Feedback on Cognitive Load

The result showed that the cognitive load in the VR navigation task is significantly influenced by using auditory feedback. Without auditory feedback, users only rely on visual scanning, which increases the mental workload. On the other hand, specifically classical and natural music categories represented lower cognitive loads than no-audio feedback and other audio categories. From this result, it supported our hypothesis (H1) that assistive auditory feedback can improve the cognitive load level due to providing external cues, which helps to make decisions easily during navigation tasks.

B. Cognitive Load–Balance Coupling in VR

From the result analysis, we got the key finding that a strong negative correlation exists between cognitive load and balance. In the classical and natural auditory feedback conditions, the participants' postural balance improved with lower cognitive load. This supports the dual-task interference theory, which suggests that the brain has limited attentional resources. When more attention is used for cognitive processing, less attention is available for motor control [33]. In VR environments, balance is already challenged by sensory conflict and spatial uncertainty. Therefore, reducing cognitive load through auditory feedback becomes especially important for maintaining balance.

C. Cognitive Load–Pupil Diameter Coupling in VR

Another key finding of this study is that a strong positive relationship between cognitive load and pupil diameter was analyzed. For the classical and natural auditory feedback conditions, we found that higher cognitive load was associated with increased pupil diameter. These outcomes indicate that pupil diameter reflects less attention on cognitive resources.

D. Design Implications for VR Navigation Systems

This pilot study provides a clear overview of VR interface design for navigation tasks. It reveals that naturalistic, continuous, and low-complexity sounds are more comfortable and user-friendly than highly dynamic audio in assistive auditory navigation systems. Classical and natural music categories perform better in terms of cognitive load, pupil diameter, and balance. While designing VR environments, researchers and

designers should consider such audio feedback as a fundamental accessibility feature, especially when considering older users or users with balance impairments.

VI. CONCLUSION AND FUTURE WORK

This section summarizes the main findings of the study and outlines potential directions for future research to further improve VR accessibility and user experience. This pilot study utilized an applicable approach for mitigating cognitive load, which is strongly associated with improved balance and pupil diameter, by using assistive directional auditory cues. We found that the classical and natural music categories were more effective than other categories. The addition of naturalistic audio feedback in VR environments helps users to become more efficient, safer, and more accessible, especially those who are vulnerable to cognitive overload or balance impairment. These results established that the selection of auditory cues is an important factor in designing VR systems. Moreover, the results also revealed the correlation between cognitive load, postural balance, and pupil diameter, which will motivate us to run a user study with many participants to investigate the interesting findings. The results will help validate the preliminary findings of this pilot study and provide stronger evidence for designing VR systems that incorporate effective auditory cues to improve user safety, efficiency, and accessibility.

REFERENCES

- [1] D. Kahneman and J. Beatty, "Pupil diameter and load on memory", *Science*, vol. 154, no. 3756, pp. 1583–1585, 1966.
- [2] G. H. Small, L. G. Brough, and R. R. Neptune, "The influence of cognitive load on balance control during steady-state walking", *Journal of biomechanics*, vol. 122, p. 110466, 2021.
- [3] M. R. Mahmud, A. Cordova, and J. Quarles, "Multimodal feedback methods for advancing the accessibility of immersive virtual reality for people with balance impairments due to multiple sclerosis", *IEEE Transactions on Visualization and Computer Graphics*, vol. 30, no. 11, pp. 7193–7202, 2024.
- [4] S. Barza et al., "Mozart's sonata k. 448 as a paradigm in cognitive neuroscience: Evidence and hegemony", 2025.
- [5] E. Borella, B. Carretti, M. Grassi, M. Nucci, and R. Sciore, "Are age-related differences between young and older adults in an affective working memory test sensitive to the music effects?", *Frontiers in Aging Neuroscience*, vol. 6, p. 298, 2014.

- [6] Y. Xing et al., “Mozart, mozart rhythm and retrograde mozart effects: Evidences from behaviours and neurobiology bases”, *Scientific reports*, vol. 6, no. 1, p. 18744, 2016.
- [7] N. Jaušovec and K. Habe, “The influence of mozart’s sonata k. 448 on brain activity during the performance of spatial rotation and numerical tasks”, *Brain topography*, vol. 17, no. 4, pp. 207–218, 2005.
- [8] Y. Shen, “The effect of soft background music without lyrics on attention level of high school students”, in *2023 2nd International Conference on Social Sciences and Humanities and Arts (SSHA 2023)*, Atlantis Press, 2023, pp. 972–979.
- [9] M. Mohammadian, R. Kazemi, S. Mollahoseini, and A. Smith, “The effect of music on working memory performance in noisy simulated open-plan offices: An experimental study”, *Journal of Health Sciences and Surveillance System*, vol. 13, no. 2, pp. 201–208, 2025.
- [10] J. J. Alvarsson, S. Wiens, and M. E. Nilsson, “Stress recovery during exposure to nature sound and environmental noise”, *International journal of environmental research and public health*, vol. 7, no. 3, pp. 1036–1046, 2010.
- [11] O. Medvedev, D. Shepherd, and M. J. Hautus, “The restorative potential of soundscapes: A physiological investigation”, *Applied Acoustics*, vol. 96, pp. 20–26, 2015.
- [12] K. Ellahiyoun et al., “Effects of task difficulty and music expertise in virtual reality: Observations of cognitive load and task accuracy in a rhythm exergame”, *arXiv preprint arXiv:2507.06691*, 2025.
- [13] S. K. Helps, S. Bamford, E. J. Sonuga-Barke, and G. B. Söderlund, “Different effects of adding white noise on cognitive performance of sub-, normal and super-attentive school children”, *PloS one*, vol. 9, no. 11, e112768, 2014.
- [14] S. C. Peres et al., “Auditory interfaces”, *HCI beyond the GUI: design for haptic, speech, olfactory, and other nontraditional interfaces*, pp. 147–195, 2008.
- [15] R. Skarbez, F. P. Brooks Jr, and M. C. Whitton, “A survey of presence and related concepts”, *ACM computing surveys (CSUR)*, vol. 50, no. 6, pp. 1–39, 2017.
- [16] M. M. Hossain and M. R. Mahmud, “Too much to sense? investigating cognitive load and fatigue challenges due to multisensory feedback in immersive environments”, in *2025 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, IEEE, 2025, pp. 131–138.
- [17] M. R. Mahmud, A. Cordova, and J. Quarles, “Auditory, vibrotactile, or visual? investigating the effective feedback modalities to improve standing balance in immersive virtual reality for people with balance impairments due to type 2 diabetes”, in *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, IEEE, 2023, pp. 573–582.
- [18] B. J. Brickman, L. J. Hettinger, and M. W. Haas, “Multisensory interface design for complex task domains: Replacing information overload with meaning in tactical crew stations”, *The International Journal of Aviation Psychology*, vol. 10, no. 3, pp. 273–290, 2000.
- [19] M. R. Mahmud, M. Stewart, A. Cordova, and J. Quarles, “Auditory feedback for standing balance improvement in virtual reality”, in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE, 2022, pp. 782–791.
- [20] A. Scharine, K. Cave, and T. Letowski, “Auditory perception and cognitive performance”, *Helmet-mounted displays: Sensation, perception, and cognition issues*, pp. 391–489, 2009.
- [21] F. S. Ricci, L. Liguori, E. Palermo, J.-R. Rizzo, and M. Porfiri, “Navigation training for persons with visual disability through multisensory assistive technology: Mixed methods experimental study”, *JMIR Rehabilitation and Assistive Technologies*, vol. 11, no. 1, e55776, 2024.
- [22] C. Chai et al., “The effects of various auditory takeover requests: A simulated driving study considering the modality of non-driving-related tasks”, *Applied Ergonomics*, vol. 118, p. 104252, 2024.
- [23] H. F. B. Vargas, *Auditory Feedback in Integrated Development Environments to Provide Cognitive Support for Novice Developers*. Nova Southeastern University, 2021.
- [24] M. R. Mahmud, M. Stewart, A. Cordova, and J. Quarles, “Vibrotactile feedback to make real walking in virtual reality more accessible for people with and without mobility impairments”, in *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology*, 2025, pp. 1–2.
- [25] M. R. Mahmud, M. Stewart, A. Cordova, and J. Quarles, “Vibrotactile feedback to make real walking in virtual reality more accessible”, *arXiv preprint arXiv:2208.02403*, 2022.
- [26] M. R. Mahmud, M. Stewart, A. Cordova, and J. Quarles, “Standing balance improvement using vibrotactile feedback in virtual reality”, in *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, 2022, pp. 1–11.
- [27] M. R. Mahmud, A. Cordova, and J. Quarles, “Visual cues for a steadier you: Visual feedback methods improved standing balance in virtual reality for people with balance impairments”, *IEEE transactions on visualization and computer graphics*, vol. 29, no. 11, pp. 4666–4675, 2023.
- [28] M. R. Mahmud, A. Cordova, and J. Quarles, “The eyes have it: Visual feedback methods to make walking in immersive virtual reality more accessible for people with mobility impairments while utilizing head-mounted displays”, in *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*, 2023, pp. 1–10.
- [29] F. G. Paas and J. J. Van Merriënboer, “Instructional control of cognitive load in the training of complex cognitive tasks”, *Educational psychology review*, vol. 6, no. 4, pp. 351–371, 1994.
- [30] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness”, *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.
- [31] S. Schepens, A. Goldberg, and M. Wallace, “The short version of the activities-specific balance confidence (abc) scale: Its validity, reliability, and relationship to balance impairment and falls in older adults”, *Archives of gerontology and geriatrics*, vol. 51, no. 1, pp. 9–12, 2010.
- [32] J. N. Setu et al., “Mazed and confused: A dataset of cyber-sickness, working memory, mental load, physical load, and attention during a real walking task in vr”, in *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, IEEE, 2024, pp. 1048–1057.
- [33] M. Woollacott and A. Shumway-Cook, “Attention and the control of posture and gait: A review of an emerging area of research”, *Gait & posture*, vol. 16, no. 1, pp. 1–14, 2002.

Human-Centered Innovation in Water Resources Management Education

A Co-Creation Approach to Developing VR Learning Experiences

Stella Kanatouri

Institute of Data Science
German Aerospace Center (DLR)
Jena, Germany
e-mail: stella.kanatouri@dlr.de

Heribert Nacken

Academic and Research Department Engineering
Hydrology
RWTH Aachen University
Aachen, Germany
e-mail: nacken@lfi.rwth-aachen.de

Sina Truckenbrodt

Institute of Data Science
German Aerospace Center (DLR)
Jena, Germany
e-mail: Sina.Truckenbrodt@dlr.de

Friederike Klan

Institute of Data Science
German Aerospace Center (DLR)
Jena, Germany
e-mail: Friederike.Klan@dlr.de

Abstract—Virtual Reality (VR) adoption in higher education is hindered by development costs and producing curriculum-aligned educational contents that meet learners’ needs. Whilst co-creating VR educational experiences with learners can mitigate these challenges - as it ensures content relevance and reduces development time and costs - involving learners as co-designers is challenging, especially if they lack coding expertise. This study contributes a practice-based approach to the co-creation of immersive learning experiences, that involved a design thinking workshop with learners, teaching staff and developers. The findings illustrate how a layered prototyping approach can support co-creation participants in progressing from scenario content design with detailed storyboards to immersive prototypes. The study concludes that integrating no-code immersive tools and generative Artificial Intelligence (AI) into the co-creation workflow can make VR pre-production accessible, efficient and shaped by learners’ needs – opening up potential for greater VR adoption in higher education.

Keywords-Virtual Reality; co-creation; Design Thinking; no-code prototyping; Artificial Intelligence.

I. INTRODUCTION

Virtual Reality (VR) has emerged as a transformative technology with applications that span education, healthcare, engineering, and entertainment [1]. Whilst its potential to offer immersive, experiential and collaborative learning in higher education is gaining recognition, VR adoption in higher education is constrained by development costs and limited availability of curriculum-aligned content [2]. Co-creation approaches hold a promise to mitigate these challenges, as they ensure that bespoke content development will meet learner needs and curricular goals, saving VR development time and cost later in the development process [3]. At the same time, end-user involvement in co-creating VR learning is not straightforward, particularly due to learners’ potentially limited VR expertise.

In this article, we propose a co-creation process that involves learners throughout defining learning contents, designing and developing a VR learning scenario. Measures are adopted to make VR accessible to learners and empower them to contribute meaningfully in the co-creation. In Section II, we review relevant co-creation literature in the contexts of education, software development and VR learning. These perspectives inform the methodological design of our study, described in Section III. In Section IV we present the findings, followed by a discussion in Section V. We conclude with lessons learned and implications for future work in Section VI.

II. RELATED WORKS

In this section, we review existing literature on co-creation in education, software development and VR learning contexts.

A. Co-creation in educational contexts

The teacher-centered paradigm has been challenged by learner-centered pedagogy that emphasizes among other, active engagement, autonomy, and focus on learners’ individual needs [4]. Learner-centered approaches are rooted in constructivist, emancipatory and experiential learning theories, from Dewey and Freire to Vygotsky and Piaget. They propose that learners develop deeper understanding and more critical thinking skills, when they actively shape their learning processes and when learning contents are adapted to their needs [5].

Extending from the learner-centered pedagogy paradigm, more recent research has increasingly emphasized the value of co-creating learning experiences with learners [6][7]. Co-creation advances learner-centricity to another level that goes beyond merely adapting to learners’ needs to engaging them meaningfully as equal partners in the design of learning contents and experiences [8][9]. Reported benefits of co-creation in educational contexts include increased motivation and engagement, agency and collaboration [9][10]. Co-

creation has been conceptualized as a human-centered design approach that emphasizes collective creativity [11], empathetic exploration into the lived realities and challenges of those, for whom value is being created [12] and iterative development [13]. Co-creation is especially well-suited for tackling wicked, ill-defined problems – challenges that lack straightforward solutions and require context-sensitive inquiry and collaborative solving [14][15]. Despite the growing conceptual literature on co-creation in education, there is a lack of empirical studies that offer insight into the methods used to co-create learning experiences [7][16]. Only few case studies, such as Laugaland et al. [16] and Hatsikas-Schroeder et al. [17] offer such methodological detail. Insight into methodological procedures is important for improving the quality and consistency of co-creation initiatives.

B. *Co-creation in software development*

The emphasis on collaborative design in educational contexts reflects an orientation towards human-centric principles, that also guides many practices in software development. Human-centered approaches are perceived to play a key role in increasing product innovation and improving usability, user satisfaction, software quality and successful adoption [18]-[21].

Whilst inviting user feedback during the process is a recommended practice in software development, it is often limited to two stages of the process: during requirement elicitation at the outset (mainly through user interviews), and after design solutions have been pre-defined by the design team (to provide feedback in the form of usability or user experience testing) [22][23]. By limiting user involvement to these two stages, users are essentially positioned as subjects or informants rather than as equal partners in the development process [22].

The limited user involvement can be justified by the tight schedules and budgets of software development projects [19]. It might also be explained by the technical nature of software development, where end-users may lack understanding of the technology required to provide meaningful insights and express requirements [23]. Whilst human-centered approaches, such as agile methods advocate the benefits of user involvement throughout the product development lifecycle, user involvement is mainly perceived relevant for eliciting and validating requirements [23], rather than for truly co-designing software.

C. *Co-creating VR learning experiences*

The potential of VR to enrich education and training is increasingly recognized [24]. VR may offer immersive, experiential and engaging learning experiences that stimulate hands-on engagement [25]-[28]. It may facilitate the comprehension of complex or abstract concepts – for instance, in disciplines like anatomy [29]. In addition, VR allows collaborative learning by simulating physical co-presence and enabling learners to engage in joint tasks [30].

Despite the promising immersive qualities of VR that can address the need for visualization and active, hands-on learning, VR adoption in education remains limited [2][26][28][31]. In addition to the learning curve of VR and the potential discomfort or risk of experiencing motion sickness, higher education institutions face unique barriers that inhibit the adoption of immersive technologies [2][26]. Specifically, despite the growing accessibility of VR, limited university budgets inhibit adoption in academic teaching [28], not only due to hardware costs, but also due to the costs involved in purchasing VR content or developing bespoke content. Available VR content for purchase is limited in comparison to other educational resources, making it difficult to find relevant content that aligns with the university's curriculum [28]. Developing bespoke VR content ensures close alignment with the curriculum, but it requires considerable time and cost investment, and it lacks scalability, as it is unlikely to be used by several educational institutions with different specializations and curricula. Even if contents are relevant for several institutions, bespoke contents are not necessarily available for sharing as open educational resources.

Beyond these adoption challenges, VR does not automatically lead to positive learning outcomes. Its effectiveness in supporting positive learning outcomes might depend on how well-designed a VR learning application is [2][32], including interaction and immersion aspects [24][32], content relevance, positive user experience and reduced friction [24][26]. Whether purchasing existing content or developing bespoke content, a well-designed and user-friendly VR application that is closely aligned with learning needs could potentially contribute towards embracing the technology in higher education [33].

Whilst human-centered approaches may enhance content relevance and user experience, the barriers of involving end-users as co-designers is perhaps even stronger compared to 2D software: VR is a less mainstream technology, it involves a steeper learning curve and users may lack the technical expertise to participate meaningfully in the co-creation of VR prototypes [34]. Some researchers like Nguyen et al. [34] and Flobak et al. [35] have explored strategies to enable strong user participation in the co-creation of VR learning and to make the prototyping process more accessible to them. For instance, Flobak et al. [35] involved 15- to 17-year-old pupils as co-designers of a VR scenario for exposure therapy. The pupils brainstormed ideas, created a storyboard of their scenario, which they enacted in a role-play. The role play was then recorded with a 360° camera, rendered, and subsequently made accessible via Head Mounted Display (HMD). The study showed that non-developers can contribute creatively to VR content through accessible tools like 360° video.

However, the development of fully interactive VR environments requires advanced software development skills. The study by Nguyen et al. [34] proposed a co-creation process using LEGO® bricks to develop training

scenarios with police officers. The approach enabled co-creation participants with no VR design expertise to create 3D storyboards, establishing the foundation for more detailed VR scenarios to follow later on. Yet, the transfer from artefact to VR scenarios was beyond the scope of the study.

D. Research aim

The present study considers the following gaps in research and practice, as discussed previously:

- The need for deeper methodological insight into co-creating learning experiences with learners.
- The restricted user roles in software development, where users act as informants rather than as co-designers.
- The VR development challenges (pedagogical and technical) to create relevant, well-designed VR learning experiences in a time and cost-effective manner.

In response to these gaps, this study presents a co-creation process that involved learners as co-designers of a VR learning experience. It explores how to streamline early-stage development, potentially reducing development costs and time later on, and how to collaboratively create contents that genuinely meet learners’ needs and align with the academic curriculum.

III. METHOD

In this section, we describe the methodological approach employed in our study, structured along the Design Thinking process stages.

A. Methodological overview

The objective of this research was to co-create a VR learning experience about water resource management, complementing existing university teaching practices. The co-creation brought together students pursuing a Master of Science (MSc) degree in ‘Sustainable Management – Water and Energy’, faculty members, and VR developers of the Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen University, Institute of Hydrology, in Germany.

In collaboration with the university, seawater desalination was selected as the overarching co-creation theme, as one of the topics covered by the study program. The students were presented with an open-ended design challenge: *What should a VR scenario look like that students can use to acquire the specialist knowledge and skills they need to plan a seawater desalination plant?* The design challenge of the VR learning scenario was intentionally left open for the students to define together more specifically during the co-creation. Given the complexity and open-ended nature of the design challenge - a “wicked” problem with multiple possible solutions - the design thinking approach [36] was selected to structure the co-creation.

The co-creation involved exploratory interviews and a two-day workshop. The interview discussion guide and the workshop structure are provided as supplementary material on the online platform Zenodo [37]. As the students that participated in the workshop had no previous experience with VR, the *Empower* stage was integrated into the design thinking process to familiarize them with the technology. Table I presents the design overview, illustrating the key stages and steps carried out during the process. These are described in detail in the following sections.

B. Empathize Stage

In line with design thinking, initial emphasis was placed on problem exploration and understanding students’ needs. Individual interviews were conducted with three university professors and seventeen MSc students across seven German universities. Professors were identified via word-of-mouth and students were recruited via LinkedIn and selected based on their enrollment in MSc level programs with a water resource management focus. The sample consisted of seven female and ten male students. The students’ age range 23-35 years. Interviews varied in format (telephone or video call) and duration (30-60 minutes). With the interviews participants’ perspectives of current teaching methods at their university and their associated challenges for teaching / learning were explored.

TABLE I. DESIGN OVERVIEW – DESIGN THINKING STAGES AND KEY PROCESS STEPS







Stage	EMPATHIZE 	EMPOWER 	DEFINE 	IDEATE 	PROTOTYPING 	TEST 
Methods/Tools and Output	<ul style="list-style-type: none"> • Interviews with university professors and students • Thematic area alignment with professor (curriculum fit) • Facilitator immersion into topic • Empathy maps • Challenges and needs analysis 	<ul style="list-style-type: none"> • VR introductory session to familiarize with the medium 	<ul style="list-style-type: none"> • Synthesis of interview findings • Group mind map to identify key knowledge gaps 	<ul style="list-style-type: none"> • Brainstorming ideas to address content needs and clustering 	<ul style="list-style-type: none"> • User stories • LEGO® Serious Play® 3D storyboards • VR prototypes 	<ul style="list-style-type: none"> • VR prototype testing with students
	Interviews	Co-creation workshop				

TABLE II. INTERVIEW SAMPLE WITH PROFESSORS AND STUDENTS IN THE FIELD OF ENGINEERING HYDROLOGY

Role	n (total: 20)	Type of higher education institution	Interview format	Interview duration
Professor	1	University of Applied Sciences	Telephone	30min
Professor	1	University	Telephone	30min
Professor	1	Technical University	Video call	30min
Students	3	University of Applied Sciences	Video call	60min
Students	14	Technical University	Video call	60min

Table II provides an overview of the interview sample, including participants' type of higher education institution, interview format and duration.

The key interview insights were synthesized into empathy maps, capturing students' learning styles and into a summary of their pain points and needs as learners. To deepen her contextual understanding, the researcher immersed herself in the subject during this research stage, by attending three online lectures given at RWTH Aachen University and studying the university's massive open online courses about seawater desalination.

C. Empower Stage

Following the interviews, a two-day design thinking workshop was held at RWTH Aachen University to collaboratively conceive the VR learning experience. The workshop brought together MSc students, teaching staff, and VR developers of the RWTH Aachen University. A total of 13 participants attended: eight MSc students (four female, four male), all of which had already been interviewed, two faculty members, and three VR developers. While the faculty members and developers accompanied the workshop, students acted as the main designers, contributing their perspectives as learners and shaping the goals and contents of the learning experience.

Notably, 7 of 8 students had no prior experience with VR. To support meaningful participation, the workshop began with an introduction to the task and an overview of the design thinking principles guiding the process. Students were given a 2.5-hour hands-on VR session to familiarize themselves with the medium. The VR team provided a tutorial on headset controls and setup, followed by guided exploration of two VR applications. This session allowed students to experience key VR affordances - such as spatial presence, depth, spatial audio, scale, embodied interaction, and multiplayer functionality and to familiarize themselves with navigating VR spaces.

The VR session was followed by a facilitated discussion during which, students reflected on their impressions of VR, its potential for learning, and its limitations. This initial exposure to VR was designed to empower students to develop an awareness of the potential and barriers of VR and to engage confidently in the subsequent ideation and prototyping phases, despite their lack of technical expertise.

D. Define Stage

The workshop continued with a collaborative mapping of the eight students' existing knowledge about seawater desalination. A group mind map exercise was used to surface prior understanding and identify gaps and was followed by a group discussion to reflect on missing knowledge and

learning needs. These materials were introduced into the workshop to support reflection and deepen understanding of the broader learning context. The interview insights (general learning needs) and the mind map (specific knowledge gaps) were used to re-define the design challenge.

E. Ideate Stage

The ideation session focused on generating ideas to address students' specific learning needs, whilst also considering the types of learning activities that the VR environment could enable. The ideation was a structured activity to encourage individual reflection and rapid idea generation. Participants were given Post-it® notes and a few minutes to write down as many ideas as possible. Students were split into two breakout groups. Each group was facilitated by a faculty member, who was tasked with clustering students' ideas. The clusters from the two groups were merged and refined into overarching themes.

F. Prototype Stage

To manage the complexity of designing a VR learning experience, a layered prototyping strategy was employed. This approach separated *learning content*, *narrative storyboards*, and *VR implementation layers*, allowing students to contribute meaningfully to content and structure. The three layers were implemented to progressively concretize ideas and build a detailed learning scenario that reflected the group's needs. This strategy is in line with Klafki's didactical model, which posits that didactic goals precede methodological implementations [38].

The content prototyping layer involved the development of user stories. Students worked on two user story templates, using pen and paper: the first required to reflect and elaborate on their proposed topic, learning objective, and expected learning benefit. The second asked students to formulate their concept of the same topic in a single, clear sentence: "*As a learner, I want to learn about...* (learning topic), *because...* (expected benefit)". The stories were presented and discussed in the group, establishing a foundation for framing the learning scenarios.

The narrative prototyping layer was developed using the LEGO® Serious Play® Method [39]. With LEGO® Serious Play®, students were encouraged to move beyond the reflections that were previously expressed on paper to more tangible prototypes that they built with their hands. The process started with a skills building phase that was used to get students accustomed to thinking with their hands, building using metaphors and storytelling. This phase gave learners confidence in representing their ideas without overthinking how their models look, removed barriers to

building prototypes of complex or abstract concepts (for instance, energy or sustainability) using symbolic thinking, and introduced *focused* storytelling – sharing their ideas by explaining the symbolic meaning of each brick. The process continued with the building phase, where students had the opportunity to build individual prototypes of learning scenarios, followed by building a collective learning scenario. To build the collective model, each student proposed elements from their individual prototypes to be included. This process involved negotiating and voting to ensure consensus. The resulting collective model was a 3D storyboard, where students presented their scenes in detail.

The VR implementation prototyping layer involved the transfer from analogue to VR prototyping with ShapesXR [40], a VR tool, designed for simple 3D prototyping and real-time co-creation. The RWTH Aachen University VR developers had received an online training ahead of the workshop to familiarize themselves with using the tool. To establish a workable starting point, the VR team selected one of the individual scenario ideas and produced a preliminary mock up with ShapesXR during the workshop. The entire collaborative LEGO® scenario will be developed in VR at a later stage. Following the workshop, the researcher continued developing the prototype in VR using ShapesXR. As a novice in VR production and without coding expertise, she relied on Artificial Intelligence- (AI-) based guidance using Microsoft CoPilot [41] for step-by-step instruction on sketching the desalination plant and creating animations.

G. Test

The complete ShapesXR prototype will be tested with MSc students at RWTH Aachen University in spring 2026. The testing phase will involve structured feedback collection through observation, reflection and discussion sessions. Students' feedback will be synthesized and reported to the VR development team to inform the next stage of technical implementation. Further iterations of the VR experience will be conducted once development begins, ensuring that the design remains aligned with students' learning needs and expectations.

IV. RESULTS

A. Interview insights into learners' needs

The co-creation process led to defining the learning goals and contents, and detailing scenarios of the VR environment. The exploratory interviews uncovered students' and professors' challenges and needs about teaching and learning water resource management-related study topics.

Professors felt that students appear disengaged during lectures, with limited interest in active participation or deeper understanding. Opportunities for hands-on learning were described as rather restricted, primarily due to the university's financial constraints. For instance, the high cost of laboratory equipment prevents students from conducting

experiments themselves. Instead, a lab technician typically performs the experiments, while students observe passively. Similarly, site visits and field trips are rare, as the university's limited budget does not allow for regular excursions that could enrich practical understanding.

Student interviews echoed professors' views and revealed further challenges. Similar to professors, students reported feeling disengaged during lectures. They noted a strong emphasis on theoretical content and memorization, often without sufficient visualization or practical application. They felt that they lack opportunities to visualize complex systems, collaborate with peers, and especially practice decision-making. The rapid pace of instruction was generally seen as compromising deep understanding of the study materials, with success mainly defined by exam performance. To pass, students must memorize large volumes of content, much of which, is quickly forgotten after the assessment. A strong need for more active and experiential learning opportunities to complement theory was voiced by all students. As one of them explained: *"Theoretical lectures alone are not helpful. Practical exercises and implementation in the form of field trips, field measurements, lab experimentation are essential for us. Visualization is so important for us engineers. Equations are not just data – it's good to visualize information to make sense of it!"*

The summarized interview insights were presented in the workshop, allowing participants to develop a common awareness of their expectations of the VR learning environment, addressing the gaps in current teaching formats. After gaining first experiences with VR, students noted how their needs and expectations for active, visual, and collaborative learning aligned with the affordances of the medium, reinforcing their support for its use.

B. Mapping out knowledge and learning interests

Having explored students' pain points and needs during the interviews, the workshop started by focusing on mapping out their existing knowledge about seawater desalination, gauging specific learning topics they were interested to deep-dive into. Students already possessed theoretical knowledge about different desalination methods, energy and brine management methods, as well as about environmental, societal and economic issues that are raised due to desalinating water. Mapping out their knowledge, helped to define the key gap that might be addressed in the VR space: moving from fragmented theoretical knowledge about various aspects of desalination towards synthesizing and applying this knowledge to design an efficient, cost-effective, and at the same time sustainable seawater desalination plant. Planning a desalination plant involves managing trade-offs between energy use, environmental impact, cost and technical constraints. Even though students

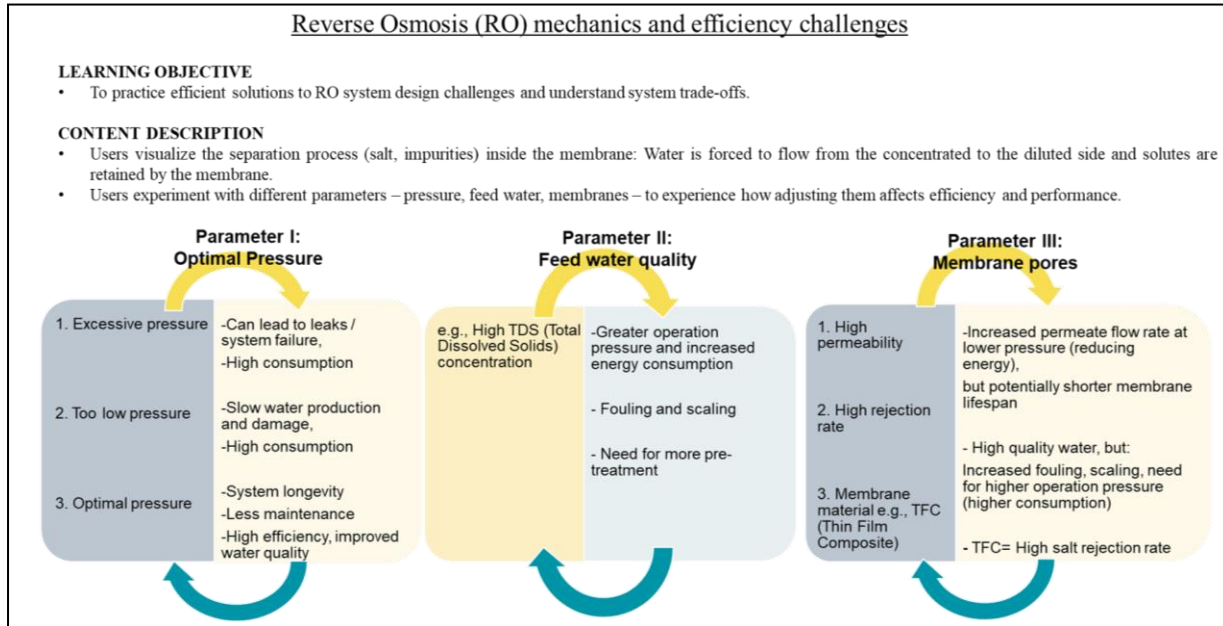


Figure 1. Exemplary user story for VR prototyping

know how a desalination plant works, they lack experience in applying their knowledge to developing innovative, sustainable and efficient solutions. Based on these insights, the design challenge evolved: *“How might we design a VR learning environment about seawater desalination that allows us to actively learn and gain experience in making various decisions needed, when planning a sustainable, cost-effective and efficient desalination plant?”* Having defined the design challenge, multiple ideas were brainstormed. These were clustered under two major themes: selecting a method (desalination method, energy and brine disposal methods or combining multiple methods) and experiencing the trade-offs of their decisions (in terms of cost-effectiveness, efficiency and sustainability).

C. Prototyping – from user stories to LEGO® models and rapid VR prototypes

User story development allowed students to assemble and expand on the rapid ideas generated during the earlier brainstorming session into coherent possible solutions. It supported reflective thinking on the rationale and the contents of the learning experience and it ensured that all voices and diverse learner insights were considered. A total of eight user stories were developed and pitched. Figure 1 presents an example of one of the user stories.

During the LEGO® Serious Play process, the students created learning scenario prototypes, first individually and then collaboratively. The collective scenario is summarized in Figure 2.

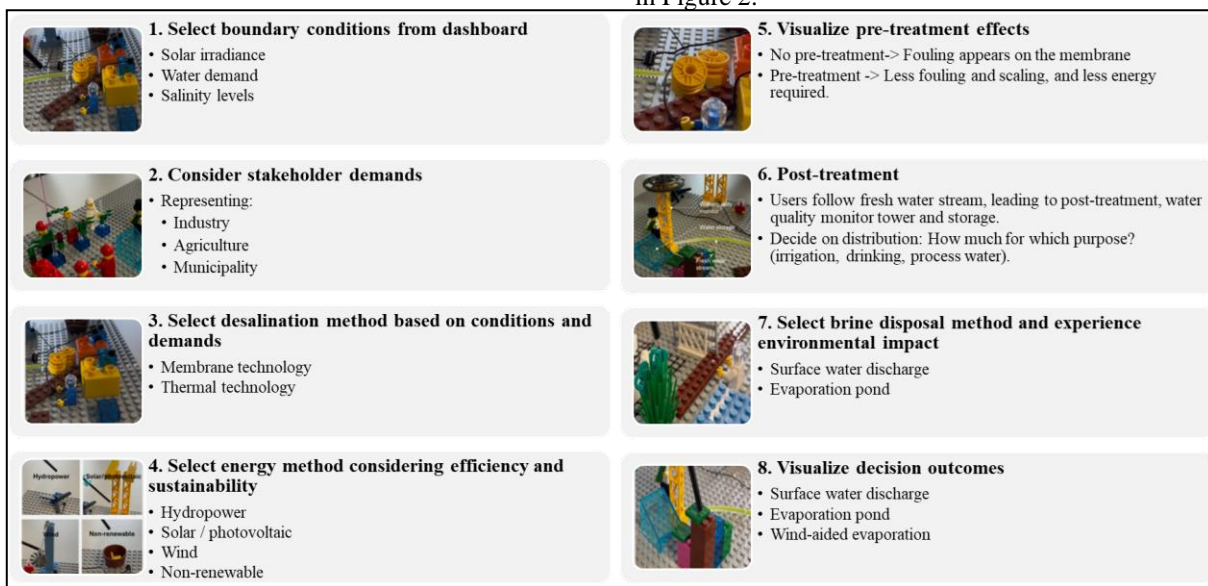


Figure 2. Collaborative storyboard of the VR scenario

The students felt that this scenario bridges their theoretical knowledge with the real-life constraints and situations they are likely to face in their future workplace. Beyond technical skills and engineering knowledge, further skills were perceived essential to effectively tackle complex engineering projects, including project management and social skills, involved when planning a desalination plant. For example, an understanding is needed in terms of how to manage different stakeholders' requirements and expectations, whilst considering specific conditions, such as water demand and cost. Moreover, by making different decisions in terms of desalination method, pre-/post-treatment, energy and brine disposal methods, learners would experience the impact of their decisions on the plant's efficiency, cost-effectiveness and sustainable management.

The transfer to VR prototyping was challenging, given the time constraints of the workshop. The complexity of the collaborative scenario, required a more manageable entry point. Hence, the VR team began by creating rough initial VR prototypes of one of the students' individual scenario ideas (Fig. 1) during the workshop, as shown in Figure 3.

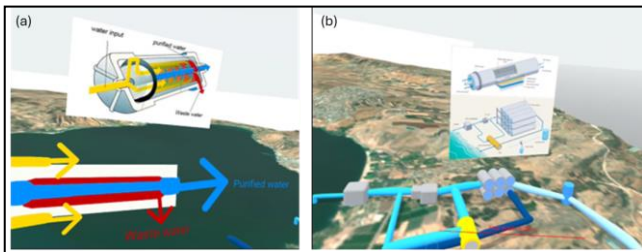


Figure 3. Impressions from the VR mock up during the workshop, simulating the salt separation process (a) and the placement of the Reverse Osmosis racks (b).

The selected learning scenario centered on visualizing the effects of adjusting pressure on water production, quality and energy efficiency in seawater desalination. It was developed further by the researcher after the workshop (Fig. 4).

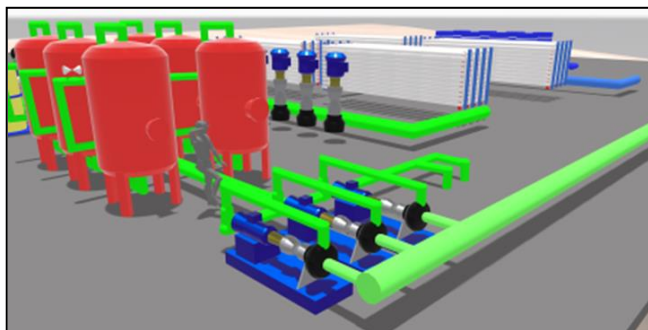


Figure 4. Impression of the VR prototype after the workshop, showing the water intake pumps and pre-treatment filters of the desalination plant.

AI provided step by step instructions for sketching components of the plant and it demonstrated how specific models, interactions and effects could look like in ShapesXR (Fig. 5).

V. DISCUSSION

The co-creation process resulted in detailed, collaboratively developed learning scenarios that aligned with the university's curriculum and were tailored to learners' needs, as students were actively involved throughout the design process. Given that the design challenge was initially ill-defined and represented a wicked, complex problem, the design thinking approach enabled a thorough exploration of the problem space.

To minimize expert-learner dynamic effects during the co-creation that could have inhibited students' voices, faculty members played a peripheral role in the workshop, whilst students were the main actors throughout the process. Two key methodological steps aimed to empower students as true co-designers:

The first step was the VR onboarding session, which helped students familiarize with the VR medium and experience its special affordances, before asking them to develop a VR scenario for a medium they do not know. On the other hand, we acknowledge that this approach had limitations: The VR session shaped students' perceptions of VR and may have influenced their ideas about a VR learning scenario. A different or more extensive sensitization process might have generated alternative ideas and scenarios. Had time not been a constraint, a broader range of VR applications could have been demonstrated, including more industry-relevant applications. Students' experiences with the medium could have also been enriched, by understanding various types of VR content and the differences between them (3D modeling, 360° content, Gaussian splat etc.) as well as different types of possible educational VR formats, such as games, virtual tours, or experiments. A more extended sensitization process may have also raised students' awareness of the technical constraints and effort involved in software development. This limitation led to the development of a complex learning scenario that complicates VR production. Moreover, the aim of involving students as equal designers was not fully achieved just by the VR onboarding. Students remain non-experts in VR and hence, were not fully capable of taking technical decisions about VR aspects.

The second step taken to enable students' meaningful involvement in the co-creation, was the layered prototyping process (content, narrative, VR implementation prototypes), which allowed ideas and prototypes to be developed stepwise, rather than immersing learners in VR prototyping immediately - a step that might have distracted from reflecting on learning objectives and content before considering aspects, inherent to VR. Instead, prototyping began with tools that students are familiar with - pen and paper. User story development allowed individual reflection and elaboration on first ideas. These stories became more tangible when building LEGO® prototypes. The LEGO® process unleashed creativity, as it guided students to work with metaphors and symbolism to represent even highly abstract concepts (e.g., sustainability). This resulted in the development of multiple, tangible and detailed learning

How do I model a high-pressure pump?

Take me through making on of those with shapes XR

Absolutely! Let's build a **high-pressure centrifugal pump** for a **seawater desalination plant** in **Shapes XR**, step by step.

Step 1: Pump Body (Main Structure)

- **Shape:** Large Cylinder
- **How to Use:** This will form the **core of the pump**, where seawater enters and gets pressurized.
- **Positioning:** Place it horizontally for a centrifugal pump.

Step 2: Motor Section

- **Shape:** Smaller Cylinder
- **How to Use:** Attach this to one side of the main pump body to represent the **motor housing**.
- **Detailing:** Use **thin cubes** to add bolts and labels.

Step 3: Inlet & Outlet Pipes

- **Shape:** Long Thin Cylinders
- **How to Use:** Attach one **cylinder** to the front (seawater intake) and another to the back (pressurized output).
- **Positioning:** Angle them slightly to simulate flow direction.

Can you show me what a high pressure pump would look like in shapes xr if I followed these instructions?




Figure 5. AI-based instruction [41] on modeling a high-pressure pump in ShapesXR and a visualization of the result of following these instructions.

scenarios in a short time. The transition to VR mock ups during the workshop was a necessary step, as analogue prototypes might deviate from the final VR product, missing implementation aspects, specific to this medium. The real-time VR prototyping in the workshop had limited success, mainly due to time constraints. Whilst ShapesXR is a tool for real-time co-creation, more time and extensive experience with the tool was essential to build testable VR prototypes in real-time. Further work on the VR prototypes after the workshop resulted in more complete versions, showing that familiarity with the tool is needed for the creation of test-ready prototypes.

VI. CONCLUSION AND FUTURE WORK

Extending the VR prototyping experience after the workshop provided valuable insight and learnings for future studies. One key learning is that, rapid VR prototyping tools like ShapesXR, can offer a pre-production workflow, which results in a potentially more streamlined process, with ultimately lower development costs and accelerated iteration cycles. Production is postponed to a later stage, once the VR prototypes have been refined to closely match needs and user experience.

Secondly, as ShapesXR does not require coding expertise and includes easy-to-learn sketching tools, it is accessible to non-experts. AI can be used to guide the entire prototyping process, lowering barriers for non-experts even further. It is anticipated that integrating no-code, VR prototyping tools and generative AI into the co-creation process could empower educational institutions to prototype their own VR learning experiences in the future. This shift towards no-code immersive design tools combined with AI could democratize immersive education.

Emerging developments are likely to improve accessibility and lower design barriers for non-VR experts

even further. For instance, beyond 3D sketching in ShapesXR, the tool allows to import external 3D models, which typically need to be purchased. Generative AI is now close to overcoming this cost barrier, as it allows the generation of 3D models from 2D images - the quality of the 3D models is improving. Moreover, developments in the field are now also enabling the automatic generation of entire, navigable, 3D worlds from a single 2D image [42]. Figure 6 presents 3D desalination plants for users to navigate through, created from 2D images and by using a simple text prompt.



Figure 6. Navigable 3D seawater desalination plant generated by Marble Worlds AI [42].

The potential of leveraging generative AI for VR content creation is supported by recent literature. For instance, Vallasciani et al. [43] presented a user study, where participants created a VR virtual museum experience with the use of generative AI. Linares-Pellicer et al. [44] highlighted that generative AI can address the bottleneck in VR educational content creation, as faculty members can use it from initial idea generation to programming and creating high-quality immersive experiences without technical expertise.

Concluding, our study has proposed a co-creation methodology to empower learners as co-designers and has offered lessons about streamlining the VR production process with no-code immersive prototyping tools. We envision that building no-code VR prototypes with AI assistance - from asset creation, to world-building - will increasingly open up possibilities for learners and lecturers to create their own immersive educational contents, tailored to their needs.

ACKNOWLEDGMENT

This contribution was made possible thanks to the generous support of the RWTH Aachen University, Institute of Hydrology, especially the students, faculty members, and VR team that participated in this study. Funding by the German Federal Ministry of Research, Technology and Space via the research project “ThWIC: Experimentelle Lernumgebung für Wasserwirtschaft in sozial gemischter Realität (WaterLab)” (grant: 03ZU1214JB) as part of the initiative “Clusters4Future” is gratefully acknowledged.

REFERENCES

- [1] CreativeNews.io: *Virtual reality: evolution, applications, and implications across diverse sectors*. [Online]. Available from: <https://creativenews.io/research-reports/virtual-reality-evolution-applications-and-implications-across-diverse-sectors/> Accessed: 2026.02.25
- [2] A. D. Samala et al., “Virtual reality in education: global trends, challenges, and impacts – game changer or passing trend?”, *Discover Education*, vol. 4, article 229, pp. 1-45, Jul. 2025, doi:10.1007/s44217-025-00650-z.
- [3] A. Kwan, M. Schroeck, and J. Kawamura: *How cocreation is helping accelerate product and service innovation*. [Online]. Available from: <https://www.deloitte.com/us/en/insights/industry/manufacturing-industrial-products/industry-4-0/cocreation-accelerating-product-innovation.html> Accessed: 2026.02.25
- [4] N. Bremner, “The multiple meanings of ‘student-centred’ or ‘learner-centred’ education, and the case for a more flexible approach to defining it”, *Comparative Education*, vol. 57(2), pp. 159-186, Sep. 2020, doi:10.1080/03050068.2020.1805863.
- [5] N. Bremner, N. Sakata, and L. Cameron, “The outcomes of learner-centered pedagogy: a systematic review”, *International Journal of Educational Development*, vol. 94, pp. 1-11, Jul. 2022, doi:10.1016/j.ijedudev.2022.102649.
- [6] K. D. Könings, S. Mordang, F. Smeenk, and L. Stassen, “Learner involvement in the co-creation of teaching and learning: AMEE guide no. 138”, *Medical Teacher*, vol. 43 (8), pp. 924-936, Nov. 2020, doi:10.1080/0142159X.2020.1838464.
- [7] M. Omland et al., “Co-creation in higher education: a conceptual systematic review”, *Higher Education*, vol. 90, pp. 1017-1047, Jan. 2025, doi:10.1007/s10734-024-01364-1.
- [8] C. Bovill, “Co-creation in learning and teaching: the case for a whole-class approach in higher education”, *Higher Education*, vol. 79, pp. 1023-1037, Dec. 2019, doi:10.1007/s10734-019-00453-w.
- [9] T. Lubicz-Nawrocka and C. Bovill, “Do students experience transformation through co-creating curriculum in higher education?”, *Teaching in Higher Education*, vol. 28(7), pp. 1744-1760, May 2021, doi:10.1080/13562517.2021.1928060.
- [10] K. E. Matthews, “Engaging students as participants and partners: an argument for partnership with students in higher education research on student success”, *International Journal of Chinese Education*, vol. 7, pp. 42-64, Jan. 2018, doi:10.1163/22125868-12340089.
- [11] E. B. N. Sanders and P. J. Stappers, “Co-creation and the new landscapes of design”, *CoDesign*, vol. 4(1), pp. 5-18, Jun. 2008, doi:10.1080/15710880701875068.
- [12] J. Liedtka, R. Salzman, and D. Azer, “Design thinking for the greater good – innovation in the social sector”, New York: Columbia University Press, pp. 1-352, Sep. 2017.
- [13] C. Avila-Garzon and J. Bacca-Acosta, “Thirty years of research and methodologies in value co-creation and co-design”, *Sustainability*, vol. 16(6), article 2360, pp. 1-28, Mar. 2024, doi:10.3390/su16062360.
- [14] A. Newell and H. A. Simon, “Human problem solving”, Englewood Cliffs: Prentice-Hall, pp. 1-952, 1972.
- [15] J. Liedtka: *Darden Working Paper Series: Exploring the impact of design thinking in action*. [Online]. Available from: [Working-paper-Liedtka-Evaluating-the-Impact-of-Design-Thinking.pdf](#) Accessed: 2026.02.25
- [16] K. A. Laugaland, K. Akerjordet, C. T. Frøiland, and I. Aase, “Co-creating digital educational resources to enhance quality in student nurses’ clinical education in nursing homes: report of a co-creative process”, *Journal of Advanced Nursing*, vol. 79, pp. 3899-3912, Jul. 2023, doi:10.1111/jan.15800.
- [17] N. Hatsikas-Schroeder, T. Rüdüsühli Kunzmann, J. Amstutz, and J. Zimmermann, “Individuelle Lernprozesse begleiten – wie geht das? Ko-Kreation in der Lehrentwicklung mit Design Thinking”, *Zeitschrift für Hochschulentwicklung*, vol. 19(3), pp. 229-248, Oct. 2024, doi:10.21240/zfhe/19-03/13.
- N. Hatsikas-Schroeder, T. Rüdüsühli Kunzmann, J. Amstutz, and J. Zimmermann, “Accompanying individual learning processes – how does it work? Co-creation in the development of teaching with design thinking”, *Journal of higher education development*, vol. 19(3), pp. 229-248, Oct. 2024, doi:10.21240/zfhe/19-03/13.
- [18] F. Jesus, R. Rocha, I. Vanderlei, J. Araujo, and F. Freitas, “Design thinking evidence in software development: results from a systematic mapping”, *IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, IEEE Press, Dec. 2023, pp. 1-6, doi:10.1109/WETICE57085.2023.10477817.
- [19] O. Aromokeye, “Exploring the role of user-centered design in shaping effective IT product development processes”, *Iconic Research and Engineering Journals*, vol. 8(6), pp. 485-496, Dec. 2024.
- [20] J. J. Garrett, “The elements of user experience: user-centered design for the web and beyond”, 2nd ed., Berkley: Pearson Education, pp. 1-172, 2011.
- [21] T. J. Omonigho, “The importance of usability and user-centered design in web applications: a critical review of literature”, *PsyArXiv*, pp. 1-4, Dec. 2023, doi:10.31234/osf.io/tnu47.
- [22] J. Livari and N. Livari, “Varieties of user-centeredness: an analysis of four systems development methods”, *Information*

- Systems Journal, vol. 21(2), pp. 125-153, Feb. 2011, doi:10.1111/j.1365-2575.2010.00351.x.
- [23] Q. Omerkhel, O. M. Yusop, S. A. Ismail, and A. Azmi, "User involvement approach in agile software development: a systematic literature review", *Journal of Information System and Technology Management*, vol. 8, pp. 1-20, Sep. 2023, doi:10.35631/JISTM.832001.
- [24] X. Han, H. Luo, Z. Wang, and D. Zhang, "Using virtual reality for teacher education: a systematic review and meta-analysis of literature from 2014 to 2024", *Frontiers in Virtual Reality*, vol. 6, article 1620905, pp. 1-16, Sep. 2025, doi:10.3389/frvir.2025.1620905.
- [25] C. Christou, "Virtual reality in education" in *Affective, interactive and cognitive methods for e-learning design: creating an optimal education experience*, A. Tzanavari, and N. Tsapatoulis, Eds., Hershey: IGI Global Scientific Publishing, pp. 228-243, 2010, doi:10.4018/978-1-60566-940-3.ch012.
- [26] S. Kavanagh, A. Luxton-Reilly, B. Wuensche, and B. Plimmer, "A systematic review of virtual reality in education", *Themes in Science and Technology Education*, vol. 10(2), pp. 85-119, 2017.
- [27] F. Mallek, T. Mazhar, S. F. A. Shah, Y. Y. Ghadi, and H. Hamam, "A review on cultivating effective learning: synthesizing educational theories and virtual reality for enhanced educational experiences", *PeerJ Computer Science*, vol. 10, pp. 1-41, May 2024, doi: <https://doi.org/10.7717/peerj-cs.2000>.
- [28] S. Vats and R. Joshi, "The impact of virtual reality in education: a comprehensive research study", in *Transfer, Diffusion and Adoption of Next-Generation Digital Technologies. TDIT 2023. IFIP Advances in Information and Communication Technology*, vol. 699, S. K. Sharma, Y. K. Dwivedi, B. Metri, B. Lal, and A. Elbanna, Eds., Cham: Springer, pp. 126-136, 2023.
- [29] H. S. Maresky, et al., "Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education", *Clinical Anatomy*, vol. 32, pp. 238-243, Oct. 2018, doi: 10.1002/ca.23292.
- [30] L. Paulsen, S. Dau, and J. Davidsen, "Designing for collaborative learning in immersive virtual reality: a systematic literature review", *Virtual Reality*, vol. 28, article 63, pp. 1-17, Mar. 2024, doi:10.1007/s10055-024-00975-4.
- [31] M. Sümer and D. Vaněček, "A systematic review of virtual and augmented realities in higher education: Trends and issues", *Innovations in Education and Teaching International*, vol. 62(3), pp. 811-822, Jul. 2024, doi:10.1080/14703297.2024.2382854.
- [32] Y. Lin, G. Wang, and A. Suh, "Exploring the effects of immersive virtual reality on learning outcomes: a two-path model", in *Augmented Cognition. Human Cognition and Behavior (HCI 2020)*, D. Schmorow, and C. Fidopiastis, Eds., Cham: Springer, pp. 86-105, 2020, doi:10.1007/978-3-030-50439-7_6.
- [33] O. H. Fares, J. Aversa, S. H. Lee, and J. Jacobson, "Virtual reality: a review and a new framework for integrated adoption", *International Journal of Consumer Studies*, vol. 48(2), article e13040, pp. 1-26, Mar. 2024, doi:10.1111/ijcs.13040.
- [34] Q. Nguyen, E. Jaspert, M. Murtinger, S. Kriglstein, and M. Tscheligi, "CoLEBricks: co-designing virtual reality scenarios with generative building blocks", *Proc. of the 2024 International Conference on Advanced Visual Interfaces (AVI 2024)*, Association for Computing Machinery, Jun. 2024, article 52, pp. 1-5. doi:10.1145/3656650.3656686.
- [35] E. Flobak et al., "Participatory design of VR scenarios for exposure therapy", *Proc. of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*, Association for Computing Machinery, May 2019, Paper 569, pp. 1-12, doi:10.1145/3290605.3300799.
- [36] H. Plattner, C. Meinel, and U. Weinberg, "Design Thinking: Innovation lernen - Ideenwelten öffnen." München: mi-Wirtschaftsbuch Verlag, pp. 1-240, 2009.
- H. Plattner, C. Meinel, and U. Weinberg, "Design Thinking: Learning innovation - Opening up idea worlds." Munich: mi-Wirtschaftsbuch Publishing, pp. 1-240, 2009.
- [37] S. Kanatouri, S. Truckenbrodt, and F. Klan: *Co-creation of VR learning scenarios: Guideline for Interviews and Workshop Structure*. [Online]. Available from: <https://doi.org/10.5281/zenodo.18492729> Accessed: 2026.02.25
- [38] W. Klafki, "Das Problem der Didaktik", *Zeitschrift für Pädagogik*, 3. Beiheft, pp. 19-62, 1963.
- W. Klafki, "The problem of didactics", *Journal of Pedagogy*, vol. 3, pp. 19-62, 1963.
- [39] D. Hillmer, "PLAY! Der unverzichtbare LEGO® SERIOUS PLAY® Praxis-Guide für Workshops, Coachings und Moderation." 2nd ed., München: Hanser, pp. 1-227, 2023.
- D. Hillmer, "PLAY! The essential LEGO® SERIOUS PLAY® Practice guide for workshops, coaching and moderation." 2nd ed., Munich: Hanser, pp. 1-227, 2023.
- [40] ShapesXR: *Bringing Ideas to life in 3D*. [Online]. Available from: <https://www.shapesxr.com/> Accessed: 2026.02.25
- [41] Microsoft CoPilot: *Your AI companion*. [Online]. Available from: <https://copilot.microsoft.com/> Accessed: 2026.02.25
- [42] F. F. Li: *Marble Worlds AI*. [Online] Available from: <https://marble.worldlabs.ai/> Accessed: 2026.02.25
- [43] G. Vallasciani, L. Stacchio, P. Cascarano, and G. Marfia, "CreAIXR: fostering creativity with generative AI in XR environments", 2024 IEEE International Conference on Metaverse Computing, Networking, and Applications (MetaCom), IEEE Press, Aug. 2024, pp. 1-8, doi:10.1109/MetaCom62920.2024.00034.
- [44] J. Linares-Pellicer, J. Izquierdo-Domenech, I. Ferri-Molla, and C. Aliaga-Torro, "Breaking the bottleneck: generative AI as the solution for XR content creation in education", in *Advanced Technologies and the University of the Future. Lecture Notes in Networks and Systems*, vol. 1140, E. Vendrell Vidal, U. R. Cukierman, and M. E. Auer, Eds., Cham: Springer, pp. 9-30, 2025.