Immersive Interfaces for Engagement and Learning: Cognitive Implications

Jérôme Dinet, Muneo Kitajima

Immersive Interfaces for Engagement and Learning: Cognitive Implications

Jérôme Dinet
University of Lorraine
BP 13397, Nancy Cedex, France
jerome.dinet@univ-lorraine.fr

Munéo Kitajima
Nagaoka University of Technology
Nagaoka, Niigata, Japan
mkitajima@kjs.nagaokaut.ac.jp

ABSTRACT

Immersive Virtual Environments are distinct from other types of multimedia learning environments. But, if immersion defined the subjective impression that one is participating in a comprehensive and a realistic experience, immersive-ness is generally defined only from a systemic point of view (e.g., capacity to track users' movements, facial expressions and gestures, quality of appearance, combination of multi-sensory information, design of the virtual world). Moreover, nowadays, it does not exist a robust theoretical framework to describe and to predict immersive-ness from a user-point of view. So this paper is aiming to assume that (a) immersive-ness should be defined from a cognitive user point of view, and that (b) the cognitive architecture called MHP/RT (for Model Human Processor with Realtime Constraints) is relevant to understand and to predict immersive-ness. After a presentation of the MHP/RT model and the distributed memory system related to conscious and unconscious processes, we present the conditions necessary to produce an "immersive experience" for the user, and a case study is described as an example. Theoretical and methodological perspectives are discussed.

CCS CONCEPTS

• **Computer systems organization → Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability; • **Human-centered computing → Collaborative interaction**;

KEYWORDS

cognitive architecture, immersive interfaces

ACM Reference Format:
Jérôme Dinet and Munéo Kitajima. 2018. Immersive Interfaces for Engagement and Learning: Cognitive Implications. In *VRIC '18: Virtual Reality International Conference - Laval Virtual VRIC '18, April 4–6, 2018, LAVAL, France.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3234253.3234301

VRIC '18, April 4–6, 2018, LAVAL, France

ACM ISBN 978-1-4503-5381-6/18/04…$15.00
https://doi.org/10.1145/3234253.3234301

1 INTRODUCTION

The principal function of immersive systems is to synthesize multimodal perceptions that do not exist in the current physical environment, thus immersing users in a seamless blend of visual, aural and motor information [17]. Our present paper is based on three findings:

(1) On the one hand, some prior research has established that while the use of immersive, concrete and familiar examples can provide many important benefits for learning, it is also associated with some serious disadvantages, particularly in learners' ability to recognize and transfer their knowledge to new analogous situations (for a synthesis, see [3];
(2) On the other hand, if immersion defined the subjective impression that one is participating in a comprehensive and a realistic experience [7], immersive-ness is generally defined from a systemic point of view (e.g., "quality" of visual information, combination of auditory and visual stimuli). Moreover, several authors consider that immersion in a virtual environment is reduced by extraneous distractions and is increased by factors that facilitate direct interaction with the virtual environment and the performance of virtual environment task activities [21];
(3) If immersion (or presence, or involvement) is assessed by collecting physiological data or verbalization (e.g., by using the Presence Questionnaire elaborated by Witmer and his colleagues [21]), it does not exist a robust theoretical framework to describe and predict immersive-ness from a user-point of view. Even if some authors tried to mobilize concepts issued from psychology and ergonomics to explain immersive-ness (e.g., [6, 19]), there is no model actually to help us to better understand "when", "where" and "why" immersion exists or not.

So, in this paper we assume that (a) immersive-ness should be defined from a cognitive user point of view, and that (b) the cognitive architecture called MHP/RT (for Model Human Processor with Realtime Constraints) is relevant to understand and to predict immersive-ness.

2 PERCEPTUAL-COGNITIVE-MOTOR PROCESSES AND MEMORY

2.1 Cognitive user point of view for immersive-ness

At the 0-th order approximation, a person interacts with the ever changing environment by running an endless cycle of perceiving the external environment through five senses, i.e., taste, sight, touch, smell, and sound, via sensory neurons as parallel processing and acting to the external environment through body parts, e.g., limbs, eye balls, and so on, via motor neurons as serial processing (see Figure 1). As s/he perceives the results of the movement of his/her body parts as well as the changes of the external environment as time goes by, the next cycle of perceive–motor should occur. Interneurons in-between the sensory neurons and motor neurons convert the input patterns to the output patterns – these constitute a Perceptual-Cognitive-Motor process (PCM process).

This paper suggests that any subjective feelings should emerge while PCM processes are running in specific modes in specific external environments. Immersive feeling is one type of such feelings that can be characterized by a specific running mode of the PCM processes in a specific external environment. The purpose of this paper is to define immersive feeling as an intersection of PCM processes and external environment.

2.2 PCM as a Dual Process (Conscious and Unconscious Processes)

This section introduces a cognitive architecture, MHP/RT (Model Human Processor with Realtime Constraints) proposed by [12, 13], that is capable of simulating action selection processes in *any* situations in which human beings are interacting with the ever- changing environment including not only real physical environments but also artificial virtual environments. MHP/RT consists of memory and action selection processes, and describes in detail not only how action selections are carried out and which action will be performed but also how the results of action selections are stored in memory (see [10] for a full description of the architecture and its applications). The former process is an engagement process and the latter learning.

Action selection and memorization is a cyclic process that works endlessly while one lives in the world. As one interacts with the environment, memory is gradually structured as multi-dimensional memory frames (MD-frames); memory is represented in multi- dimensional frames because the input for memory is represented in

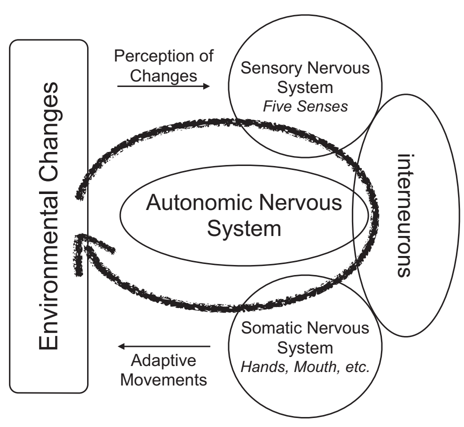

Figure 1: Continuous cyclic loop of perception and movement (Figure 1, [14]).

the dimensions of space, time, vision, audio, smell, and so on, that constitute the physical and chemical environment surrounding human beings.

Constraints on behavioral processing are imposed by *conscious* and *unconscious* processes, System 2 and System 1 of Two Minds [8, 9], respectively, and behavior must be *synchronized* with the ever-changing external and internal environments, which is a form of self-organization. As one grows while carrying out the cyclic processes of action selection and memorization, s/he develops his/her memory and shows distinct behavioral characteristics [14, 16].

The left portion of Figure 2 illustrates MHP/RT [12, 13], which consists of five autonomous sub-systems. MHP/RT is an extension of a version of dual processing theories, Two Minds, proposed by [8, 9]. Two Minds consists of unconscious processes, System 1, and conscious processes, System 2. System 1 is a fast feed-forward control process driven by the cerebellum and oriented toward immediate action. In contrast, System 2 is a very slow feedback control process driven by the cerebrum and oriented toward future action. MHP/RT focuses on synchronization between System 1 and System 2 in the information flow from the perceptual system from the environment at the left end to the motor system at the right end. This "synchronization" concerns internal synchronization. This paper, however, extends the notion of synchronization in order to deal with feeling of immersive-ness as will be described in the next sections.

Output from the perceptual system is diverted into three paths, one path leads to the conscious process of System 2, the other leads to the unconscious process of System 1, and the last one leads to the memory system. It is important to note that the path from the perceptual system to the memory system is the fastest path among the three paths, and this means that part of memory relevant to the perceived object is activated before the unconscious and conscious processes initiate to process the perceived information. Information in memory activated by the input from the environment is become available to System 1 and 2, which does have effect on whether feeling of immersive-ness should occur or not. System 1 and 2 work in synchronous with each other but the memory process works asynchronously with System 1 and 2. The dotted oval shows the process of memorization of output from the motor process.

2.3 Distributed Memory as Multidimensional frame

Figure 2 illustrates how each MD-frame is created as the result of working of autonomous processes in MHP/RT and how MD-frames are mutually interrelated. This essentially details the process "Memorization of Behavioral Actions and their Results" shown by the dotted oval in the diagram of MHP/RT by considering neuronal activities that actually happen. The basic idea is that each autonomous system has its own memory.

The five MD-frames are defines as follows:

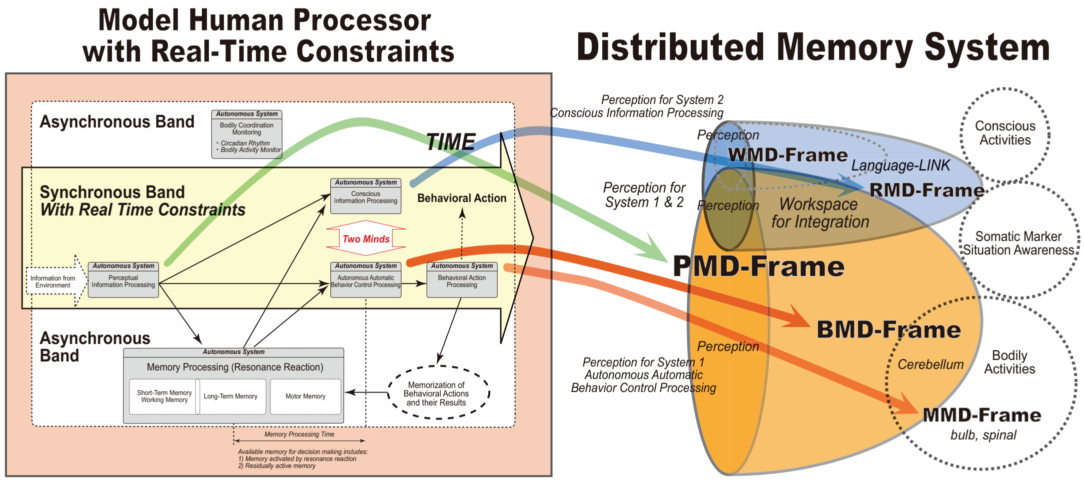

Figure 2: MHP/RT (Figure 3, [13]) and the distributed memory system.

- **PMD (Perceptual Multi-Dimensional)-frame** constitutes perceptual memory as a relational matrix structure. It collects information from external objects followed by separating it into a variety of perceptual information, and recollects the same information in the other situations, accumulating the information from the objects via a variety of different processes. PMD-frame incrementally grows as it creates memory from the input information and matches it against the past memory in parallel.
- **MMD (Motion Multi-Dimensional)-frame** constitutes behavioral memory as a matrix structure. The behavioral action processing starts when unconscious autonomous behavior shows after one's birth. It gathers a variety of perceptual information as well to connect muscles with nerves using spinals as a reflection point. In accordance with one's physical growth, it widens the range of activities the behavioral action processing can cover autonomously.
- **BMD (Behavior Multi-Dimensional)-frame** is the memory structure associated with the autonomous automatic behavior control processing. It combines a set of MMD-frames into a manipulable unit.
- **RMD (Relation Multi-Dimensional)-frame** is the memory structure associated with the conscious information processing. It combines a set of BMD-frames into a manipulable unit.
- **WMD (Word Multi-Dimensional)-frame** is the memory structure for language. It is constructed on a very simple one-dimensional array.

2.4 Functional flow structure, layered structure, and evolving cyclic network structure

Figure 3 illustrates how the MD-frames are interrelated by introducing three structures representing different view points.

- **The functional flow structure** describes memory activation paths starting from PMD-frames to WMD-frame, RMD-frame, or BMD-frame to MMD-frame, which describes how perception triggers motions. Notice that MMD-frame is the terminal MD-frame and therefore the paths from WMD-frame to RMD-frame, and RMD-frame to BMD-frame exist.
- **The layered structure** consists of the three layers:
  1) *PMD-frame–WMD-frame layer* is the top layer controlled by words. It consists of simple one-dimensional array of symbols, logically constructed language, grammars that specify language use, etc.
  2) *PMD-frame–RMD-frame layer* is the middle layer that resides on the behavioral eco-network for the individual. In this layer, one acquires the meaning of behavior in the social ecology.
  3) *PMD-frame–BMD/MMD-frame layer* is the bottom layer that creates a behavioral eco-network for the individual. This is a cyclic network starting from PMD- towards MMD-frame, and returning to PMD-frame.
- **The evolving cyclic network structure** depicted by a blue spiral refers to the fact that respective autonomous systems generate distributed memories for their use, and the memories are cyclically related and in effect topological.

These three features enable pipelining the processes for establishing synchronization in perception of the external environment, System 1, and System 2. This cyclic connection is critical to understand the relationship between behavior and memory.

3 MHP/RT'S FOUR-PROCESSING MODES: USE AND MODIFICATION OF MEMORY

As shown by Figure 4, MHP/RT works in one of four different modes when one looks at it from *a particular event* that occurred at the absolute time T. Experience associated with a person's activities is characterized by a series of events, each of which should be recognized as a person consciously. It is important to note, however, that an experience represented as a series of consciously identified events by a person has to be regarded as the results of unrecognized unconscious activities: metaphorically speaking, consciousness is one of tips of icebergs that appear above the sea level, and the tips are interrelated with each other via the unseen relationships established below the sea level. A system of icebergs develops in the natural condition of seawater and atmosphere, which may or may not be trivial for any people. In this paper, "a particular event" refers to an event that should make a person feel immersive experience.

Two are before the event in which MHP/RT uses memory for engaging in activities and the other two are after the event in which it modifies memory which results in learning of the event.

3.1 Engagement Modes: Conscious Anticipation and Unconscious Tuning "Before Event"

A person can engage in an event to happen in the future in the following two modes:

- **System 2 Before Mode:** MHP/RT consciously uses memory before the event for anticipating the future event which takes relatively long time ($T-$ some amount of time).
- **System 1 Before Mode:** MHP/RT unconsciously uses memory just before the event, say 100 milli seconds before the event for automatic preparation for the future event ($T-$ a few hundreds milli seconds).

3.2 Learning Modes: Unconscious Adaptation and Conscious Reflection "After Event"

A person can learn from the past event in the following two modes:

- **System 1 After Mode:** MHP/RT unconsciously tunes the current network connections related to the past event for better performance for the same event in the future ($T+$ a few hundreds milli seconds).
- **System 2 After Mode:** MHP/RT consciously reflects on the past event resulting in structural changes in memory ($T+$ some amount of time).

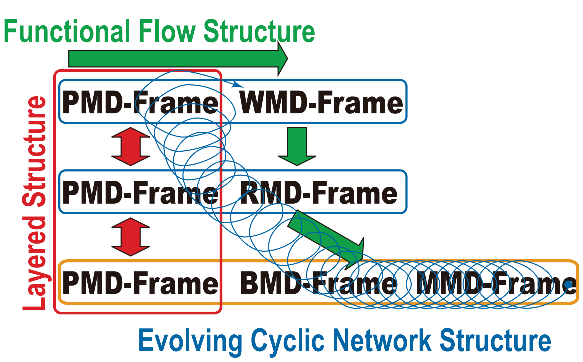

Figure 3: Functional flow structure, layered structure, and evolving cyclic network structure (Figure 4, [15]).

4 WEAK SYNCHRONIZATION WITH ENVIRONMENT FOR IMMERSIVE FEELING

This section explores how "immersive feeling" can be defined by using the theoretical constructs shown in Figures 2, 3, and 4. In order to do this, it would be useful to introduce the claim made by Damasio [2]:

> *Emotion emerges when consciousness is recognized for the first time. Feeling appears when the emotion is analyzed ecologically and recognized at the later time.*

According to this claim, immersive feeling should occur in System 2 After Mode in Figure 4, in which a person reflects on a specific event that is created through interactions with an artifact expected to provide immersive experience.

Conscious reflection on the event carried out in System 2 After Mode should result in a memory trace that integrates multiple memory frames in the region "workspace for integration" illustrated in Figure 2. It is important to note that memory activation process is a totally parallel process, and this means that there is no way of knowing which part of activated memory is used. It solely depends on which object MHP/RT is processing. MHP/RT's resonance process makes available the relevant part of activated knowledge through resonance. At the time when a person is in the System 2 After Mode, there are regions which have been activated in System 2 Before, System 1 Before and System 1 After Modes when dealing with the past event to be reflected on in System 2 After Mode. A memory trace relevant with the event is created with the feeling associated with the event.

4.1 Two Experiential Modes in Immersive Environment

This section discusses necessary conditions for a person to feel immersive-ness while interacting with an artifact.

Suppose a person is put in an environment in which s/he is expected to experience an immersive feeling. The current perceptual stimuli should activate corresponding regions in PMD-frame, which then activate relevant portions of multiple MD-frames by spreading activation through chain-firing. There are two scenarios as follows:

4.1.1 New Mode – Necessary for Feeling Immersive-ness. If there is no immersive experience associated with the current perceptual stimuli, s/he has to interact with the current environment by utilizing any portion of multiple MD-frames that have been activated by the current perceptual stimuli. Even if s/he has no previous experience of having a feeling of immersive-ness for the current perceptual stimuli, s/he would carry out a series of actions for accomplishing his/her current anticipation associated with the current perceptual stimuli.

When a series of events have happened as anticipated in terms of conscious level and unconscious level in the new environment, s/he would have carried out actions in a smooth streamlined manner, without any serious motor-level adjustment in System-1 Before Mode. As the result of the deductions, this paper suggests as follows:

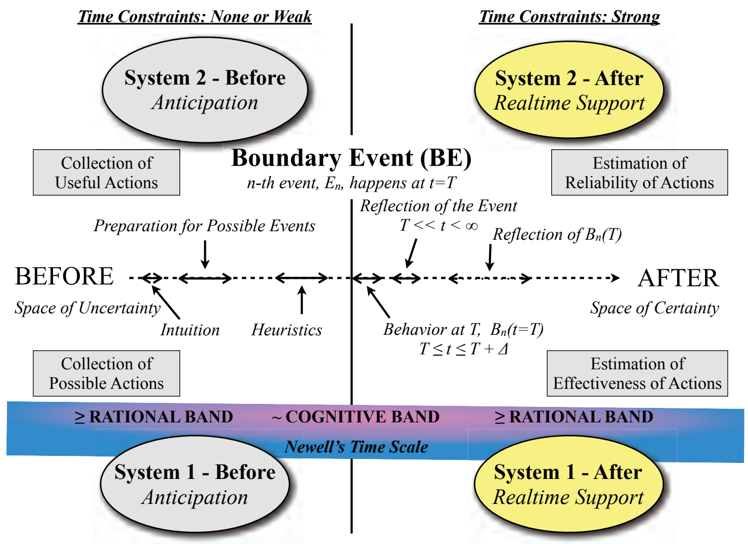

Figure 4: How the Four Processing Modes work (adapted from [11])

Immersive feeling eliciting condition for an artificial environment to have the user feel immersive-ness is 1) it must be new to him/her, 2) s/he is able to carry out actions with an anticipation activated by the artificial environment without any breakdown in performing motor-level actions (System 2 Before Mode followed by no serious adjustment required in System 1 Before Mode), 3) s/he is able to consciously recognize an event associated with the series of just-finished actions, and 4) s/he is able to reflect on the event to integrate it with the recognized feeling associated with the event (in System 2 After Mode).

An experience consists of a number of events, and those events are not independent but highly contextualized. As such, in order for an anticipation-based streamlined actions to happen, the time series of perceptual stimuli generated through the interactions between a person and the "immersive" artifact have to be well-designed for the particular person in the environment.

4.1.2 Familiar Mode – Not Necessarily Associated with Immersive Feeling. If there exists an integrated memory trace that includes the current perceptual representation as its component, it has a chance of being utilized in System 2 Before Mode for planning future actions through memory resonance processes, e.g., carrying out an action that has caused immersive feeling in the past, and/or in System 1 Before Mode to adjust the motor movements to be carried out for accomplishing the plan in such a way that they are executable at the specific environmental conditions, which happens within ~100 milli seconds time range. In Figure 3, these processes are represented as functional flow structure, starting from PMD-frame to WMD-frame and RMD-frame in System 2 Before Mode, and to the PMD–BMD–MMD-frame layer in System 1 Before Mode.

After experiencing any recognizable and distinguishable event, s/he would enter in System-1 After Mode and/or System-2 After Mode in which s/he reflects on the event to have had immersive feeling in the past resulting in strengthening the corresponding memory trace. However, the feeling recognized in reflection may not be the same as the one when s/he has experienced the event first time with the immersive-ness feeling.

4.2 "Weak Synchronization" as a necessary condition for eliciting immersive feeling

The "Immersive feeling eliciting condition" can be rephrased by using the notion of weak synchronization as follows. Normally, the term "synchronization" refers to co-occurrence of two events on two distinct streams at the same time. However, as depicted in Figure 4, a person's activity related with an event has to be considered from the four processing modes, which ranges relatively long time before and after the actual time the event happens. Therefore, this paper alternatively considers "synchronization" as the phenomena a person's activities are linked with the specific recognizable event at time T through the four processing modes, when this is satisfied, the event is called "weakly synchronized with a person's activities." If a series of events are weakly synchronized with the person's activities performed as a smooth flow of the four processing modes, s/he should have the feeling of immersive-ness.

4.3 "Strong Synchronization" as a suppressing condition for immersive feeling

The smooth flow of the four processing modes can break down when a person has to adjust his/her activity while s/he is in System-1 Before Mode in such a way that his/her movement is in synchronous with the current environment. When this happens, the condition for "weak synchronization" is not satisfied but s/he has to make efforts to reach synchronization by adjusting his/her movement. When s/he reflects on this event with strong synchronization in System-2 After Mode, he/she would have a feeling associated with anticipation-violated.

5 CASE STUDY

Pedestrian trauma represents a significant proportion of all road trauma. In particular, the safety of child pedestrians is of concern, given that a sizeable proportion of pedestrians killed and seriously injured involve children and the special value society places on its youth. At ages 6-10 years, children are at highest risk of pedestrian collision, most likely due to the beginning of independent unsupervised travel at a time when their road strategies, skills and understanding are not yet fully developed. If some simulators specially dedicated to safety of pedestrians exist, there are too expensive for schools, very large, need a high level of technical competencies and mostly used for research.

It is the reason why an immersive environment has been created by the Human Games Enterprise [1], this project being financially supported by the Grand-Nancy. This immersive environment has been conceived by taking into account the recommendations issued from prior studies directly related to immersive environment for safety of pedestrians [5, 18]. Our immersive environment is very cheap, easy to use, not based on language and adapted to the audience. In this digital environment, the user can navigate in different urban scenes, can cross streets where different kinds of vehicles and individuals can exist (Figure 5).

We conducted an exploratory study with nine young pupils (mean age = 9:5 years-old) in a French primary schools located in the West of France. Participants were volunteers for testing this innovative immersive environment dedicated to the safety of young pedestrians (Figure 6). All participants are French native speakers, none has visual impairment, and all have no experience with immersive environment before this experiment. Their cognitive abilities have been assessed by using the fifth edition of Wechsler Intelligence Scale for Children (WISC-V) [20]. No cognitive impairment has been detected. All parents agreed for the participation.

The main objectives of the exploratory study were to collect data about (i) their feelings (i.e., subjective impressions and verbalizations) and (ii) their effective behaviors. So, behaviors of the nine participants have been recorded and analyzed by several judges.

After a training test (15 minutes), each child was individually asked to navigate in different urban and dynamics scenes to go from the school to a store located in another street. So they must to search for information, to cross several streets where vehicles and other pedestrians are present, and to move. They can move by using two different modes: by "teleportation" (i.e., without motor activity) of by walk (i.e., with a motor-physical activity, with their legs).

Three main results have been obtained:

(1) it is the "teleportation" mode (vs. by walk) that has been massively chosen by the children;
(2) all the children reported having a lot of fun and saying they were "blown away" by the quality and realism of the environment;
(3) several children have adopted extremely dangerous behaviors in the immersive environment (e.g., crossing without looking, cutting a junction without even having a look at arriving cars), even though they do not have these behaviors in the reality.

This last result is the most surprising and immediately challenged us about the immersive dimension (i.e., immersive-ness) of the environments.

5.1 Discussion

The purpose of this study was not to evaluate the degree of immersiveness but the attitudes that the young participants showed towards the experimental immersive environment could be used for examining their relevance with the immersive feeling eliciting condition.

Issue 1. Even if an Immersive Virtual Environment are elaborated by using criteria defined to guarantee immersion, some users are not "immersed" while others individuals are totally "immersed"; These inter-individual differences must be take into account for the design of the Immersive Virtual Environment and ask the question of the universality of conception;

Issue 2. If immersion is often apprehended from a systemic point of view, immersive-ness must be apprehended from an end-user cognitive point of view;

Issue 3. Users who are not immersed in our Immersive Virtual Environment dedicated for young pedestrian are users for whom the condition for "weak synchronization" is not satisfied. So these users have to make efforts to reach synchronization by adjusting their movements. According to the MHP/RT perspective, the smooth flow of the four processing modes can break down when an individual has to adjust his/her activity while s/he is in System-1 Before Mode in such a way that their movements are in synchronous with the current environment.

6 CONCLUSION

From a theoretical point of view, immersion is a description of a technology, and describes the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant. But this conception is only techno-centered and cannot describe and predict human behaviors that can be observed in some case study (see our exploratory study described in previous section). Because we assume that immersive-ness is mainly a cognitive phenomenon and can be analyzed from a human-centered approach, the cognitive architecture called MHP/RT (for Model Human Processor with Realtime Constraints) is relevant to understand and to predict immersive-ness. For the future, we need to conduct investigations to better understand the reason why individuals differ in their preference for information in the various modalities provided in the Immersive Virtual Environments to enable a successful construction of their internal world models. For instance, for one person the absence of auditory information might be a crucial hindrance, whereas for another it might be hardly noticeable. And the MHP/RT can be relevant to investigate these inter-individual differences.

Figure 5: Snapshot of screens related to the immersive environment created for young pedestrian.

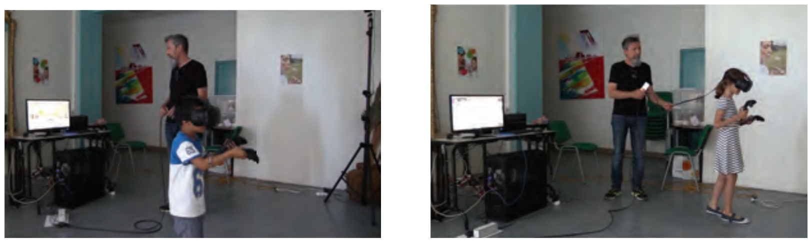

Figure 6: Two young participants of the case study.

From an applied point of view, for some authors and designers, the challenges related to Immersive Virtual Environment dedicated to learning are mainly technological (e.g., display resolution, limited tracking ability with delayed responses, providing highly localized 3-D auditory cues, haptic feedback is always limited and expensive). But, for other researchers (e.g., [4]) and us, the main problem with the use of immersive virtual environments for education is the danger of introducing new or unanticipated misconceptions due to the limited nature of the "magic" possible via this medium. For example, learners will not feel their sense of personal physical weight alter, even when the gravity field in the artificial reality they have created is set to zero [4]. The cognitive dissonance this mismatch creates, due to conflicting sensory signals, may create both physiological problems (e.g., simulator sickness) and possibly false intellectual generalizations. One part of our research is to examine the extent to which manipulating learners' visual, auditory, and tactile cues may induce subtle types of misconceptions about physical phenomena. In other words, we assume that the medium (Immersive Virtual Environment) should not detract from the learning.

This paper suggested that learning in an immersive environment takes the form of *learning a new thing*. This implies that learning associated with the feeling of immersive-ness should occur only at the first time the user encounters with the artifact designed for immersive experience. On encountering the same stimuli at later occasions, what s/he actually experiences would be better characterized as experiencing an event which s/he recognizes as the one to which s/he has experienced the feeling of immersiveness. Learning of the content itself tagged with the feeling of immersiveness should be facilitated without the feeling of immersive-ness.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 15H02784.

REFERENCES

[1] Human Games Enterprise. http://www.humangames.tv/.

[2] Antonio Damasio. 1999. *Feeling of What Happens: Body and Emotion in the Making of Consciousness.* Houghton Mifflin Harcourt, Orlando, FL.

[3] Samuel Day, Benjamin Motz, and Robert Goldstone. 2015. The Cognitive Costs of Context: The Effects of Concreteness and Immersiveness in Instructional Examples. *Frontiers in Psychology* 6 (2015).

[4] Chris Dede, Marylin C. Salzman, R. Bowen Loftin, and Debra Sprague. 1999. *Modeling and Simulation in Science and Mathematics Education.* Springer-Verlag, New York, NY, Chapter Multisensory Immersion as a Modeling Environment for Learning Complex Scientific Concepts.

[5] I. Feldstein, C. Lehsing, A. Dietrich, and K. Bengler. 2016. Pedestrian simulators for traffic research : State of the art and future of a motion lab. In *Proceedings of the 4th International Digital Human Modeling Symposium (DHM2016).*

[6] Alessandra Gorini, Claret S. Capideville, Gianluca De Leo, Fabrizia Mantovani, and Giuseppe Riva. 2011. The Role of Immersion and Narrative in Mediated Presence: The Virtual Hospital Experience. *Cyberpsychology, Behavior and Social Networking* 14, 3 (2011).

[7] Kelly S. Hale and Kay M. Stanney. 2017. *Handbook of Virtual Environments: Design, Implementation, and Applications, Second Edition.* CRC Press, Taylor and Francis Group, New York, NY.

[8] Daniel Kahneman. 2003. A Perspective on Judgment and Choice. *American Psychologist* 58, 9 (2003), 697–720.

[9] Daniel Kahneman. 2011. *Thinking, Fast and Slow.* Farrar, Straus and Giroux, New York, NY.

[10] Muneo Kitajima. 2016. *Memory and Action Selection in Human-Machine Interaction.* Wiley-ISTE. http://amazon.co.jp/o/ASIN/184821927X/

[11] Muneo Kitajima and Makoto Toyota. 2011. Four Processing Modes of *in situ* Human Behavior. In *Biologically Inspired Cognitive Architectures 2011 - Proceedings of the Second Annual Meeting of the BICA Society*, Alexei V. Samsonovich and Kamilla R. Jóhannsdóttir (Eds.). IOS Press, Amsterdam, The Netherlands, 194–199. https://doi.org/10.3233/978-1-60750-959-2-194

[12] Muneo Kitajima and Makoto Toyota. 2012. Simulating navigation behaviour based on the architecture model Model Human Processor with Real-Time Constraints (MHP/RT). *Behaviour & Information Technology* 31, 1 (2012), 41–58. https://doi.org/10.1080/0144929X.2011.602427

[13] Muneo Kitajima and Makoto Toyota. 2013. Decision-making and action selection in Two Minds: An analysis based on Model Human Processor with Realtime Constraints (MHP/RT). *Biologically Inspired Cognitive Architectures* 5 (2013), 82–93. https://doi.org/10.1016/j.bica.2013.05.003

[14] Muneo Kitajima and Makoto Toyota. 2014. Hierarchical Structure of Human Action Selections – An Update of Newell's Time Scale of Human Action. In *Procedia Computer Science, BICA 2014. 5th Annual International Conference on Biologically Inspired Cognitive Architectures*, Vol. 41. 321–325.

[15] Muneo Kitajima and Makoto Toyota. 2014. Topological Considerations of Memory Structure. In *Procedia Computer Science, BICA 2014. 5th Annual International Conference on Biologically Inspired Cognitive Architectures*, Vol. 41. 45–50.

[16] Muneo Kitajima and Makoto Toyota. 2015. Multi-dimensional memory frames and action generation in the MHP/RT cognitive architecture. In *Procedia Computer Science, BICA 2015. 6th Annual International Conference on Biologically Inspired Cognitive Architectures*, Vol. 71. 202–207.
[17] Chris Kyriakaris. 1998. Fundamental and Technological Limitations of Immersive Audio Systems. In *Proceedings of the IEEE*, Vol. 86.
[18] C. Lehsing, T. Benz, and K. Bengler. 2016. Insights into interaction. Effects of Human-Human Interaction in pedestrian crossing situations using a linked simulator environment. *IFAC – PapersOnLine* 4, 19 (2016), 138–143.
[19] Haylie L. Miller and Nicoleta L. Bugnariu. 2016. Level of Immersion in Virtual Environments Impacts the Ability to Assess and Teach Social Skills in Autism Spectrum Disorder. *Cyberpsychology, Behavior and Social Networking* 19, 4 (2016), 246–256.
[20] Wechsler. 2014. *Wechsler Intelligence Scale for Children* (5th ed.).
[21] Bob G. Witmer, Christian J. Jerome, and Mickaël J. 2005. The Factor Structure of the Presence Questionnaire. *Cyberpsychology, Behavior and Social Networking* 14, 3 (2005), 298–312.