AIVR 2024

The First International Conference on Artificial Intelligence and Immersive Virtual Reality

ISBN: 978-1-68558-154-1

April 14 - 18, 2024

Venice, Italy

AIVR 2024 Editors

Jérôme Dinet, Université de Lorraine, France
Aurelie Mailloux, Reims Hospital and University of Reims, France

AIVR 2024

Forward

The First International Conference on Artificial Intelligence and Immersive Virtual Reality (AIVR 2024), held on April 14 – 18, 2024, 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 2024 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 2024. 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 2024 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope AIVR 2024 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of AI and VR. We also hope that Venice provided a pleasant environment during the conference and everyone saved some time to enjoy this beautiful city.

AIVR 2024 Steering Committee

Muneo Kitajima, Nagaoka University of Technology (Emeritus), Japan
Manuel Caro, Universidad de Cordoba, Colombia
Carolina Cruz-Neira, University of Central Florida, USA
Aurelie Mailloux, Reims Hospital and University of Reims, France
Jerome Dinet, University of Lorraine, France

AIVR 2024

Committee

AIVR 2024 Steering Committee

Muneo Kitajima, Nagaoka University of Technology (Emeritus), Japan
Manuel Caro, Universidad de Cordoba, Colombia
Carolina Cruz-Neira, University of Central Florida, USA
Aurelie Mailloux, Reims Hospital and University of Reims, France
Jerome Dinet, University of Lorraine, France

AIVR 2024 Technical Program Committee

Suayb S. Arslan, MIT, USA
Mahajan Sagar Bhaskar, Prince Sultan University (PSU), Saudi Arabia
Concettina Buccella, University of L'Aquila, Italy
Pasquale Cascarano, University of Bologna, Italy
Courtney D. Cogburn, Columbia University, USA
Carolina Cruz-Neira, University of Central Florida, USA
Rogerio da Silva, University of Leicester, UK
Sai Anvesh Durvasula, Parabole.ai, USA
Paulo Gil, NOVA School of Science and Technology, Lisbon, Portugal
Yan Hu, Blekinge Institute of Technology (BTH), Sweden
Florentine Hüttl, University Medical Center of the Johannes Gutenberg-University Mainz, Germany
Yirui Jiang, Cranfield University, UK
Nadia Magnenat-Thalmann, University of Geneva/MIRALab, Switzerland
Sonu Mehta, Microsoft Research, India
Wolfgang Mueller, University of Education Weingarten, Germany
Moloud Nasiri, Florida Poly University, USA
Michael Nkereuwem, Federal University Dutse, Nigeria
Stamatis Papadakis, University of Crete, Greece
Vatesh Pasrija, Meta - Seattle, USA
Krzysztof Pietroszek, American University, Washington, USA
Ladislav Polák, Brno University of Technology, Czech Republic
Simant Prakoonwit, Bournemouth University, UK
Yiming Qiu, University of Michigan, USA
Lorenzo Stacchio, University of Bologna, Italy
Frank Steinicke, University of Hamburg, Germany
Davide Tebaldi, University of Modena and Reggio Emilia, Italy
Yanbo Wang, Aalborg University, Denmark
Zixuan Wang, UC San Diego, USA

Copyright Information

Table of Contents

AI for Enhancing and Preserving Dance Cultural Heritage: a Case Study on Rudolf Nureyev's Costumes

1st Silvia Garzarella
Department of the Arts
University of Bologna
Bologna, Italy
silvia.garzarella3@unibo.it

2nd Pasquale Cascarano
Department of the Arts
University of Bologna
Bologna, Italy
pasquale.cascarano2@unibo.it

3rd Lorenzo Stacchio
Department of Political Sciences,
Communication and International Relations
University of Macerata
Macerata, Italy
lorenzo.stacchio@unimc.it

***Abstract*—This paper explores the concept of costume agency in dance heritage preservation, focusing on the case study of Rudolf Nureyev (1938-1993). It considers the multifaceted role of costumes in expressing tradition, character, and performance traits, highlighting their influence on movement and identity in dance. Nureyev's career serves as an ideal case study due to his meticulous attention to costume selection and the challenges posed in preserving and valorizing his legacy. The study employs Gaussian splatting models to reconstruct 3D models of costumes from 2D images, enabling a closer examination of their design and detail. Through this interdisciplinary approach, the paper underscores the significance of costumes in conveying narrative, aesthetic, and historical aspects of dance heritage, while also showcasing the potential of technology in enhancing preservation efforts.**

***Index Terms*—Costume agency, Dance heritage, Rudolf Nureyev, Preservation, 3D reconstruction.**

I. Introduction

The cultural heritage of dance is characterized by the coexistence of tangible and intangible objects, expressing a tradition rich in extremely varied documentary traces in terms of types, dating, and geographical location [1]. Preserving such heritage entails acknowledging its plurality, which involves adopting a comparative approach to analyzing all the sources [2]. Costumes play a central role in these efforts. They not only express the visual aspects of the choreography for which they were created and the characters they represent, but also, from a technical perspective, the specific performance traits of the associated dances, those of the dancers for whom they were designed, and those of individuals who have worn them over time. In essence, the costume is a 'speaking' object, possessing an 'agency' skillfully adept at conveying crucial narratives, vital for preserving a performative event.

The notion that costumes possess agency can be linked to the "material turn" [3] of the 1980s and, more specifically, to the insights gained from studies in archaeology, anthropology, and art history, which have focused on developing methodologies to investigate the relationship between humans and objects, and the role of objects in society, thus contributing to the advancement of Material Culture Studies (MCS) [4]. By incorporating insights from various disciplines, MCS provided a nuanced perspective on the social functions, symbolic meanings, and cultural implications of objects and artifacts [5]. This multidisciplinary approach enabled researchers to explore both macro-level societal structures and micro-level individual experiences, revealing that objects not only fulfill practical functions but also symbolize social affiliations, statuses, and personal identities [6] [7]. In this respect, studies in archaeology and anthropology have been important precisely because they have highlighted more clearly how objects play a role in conditioning action and even thought [8]. Just think, returning to costumes, about how a certain type of fabric or garment can influence the movement and sensations of those who wear them, and how this aspect, can significantly influence the course of a performance. In this regard, the past decade has witnessed a significant increase in Costume Studies, initiating discussions not only on costume construction but also on their influence in shaping performances [9]. This point is important to highlight a central aspect in defining the identity of this kind of asset, which combines strong aesthetic and narrative needs with an indispensable aspect related to functionality.

In the field of dance, in which the moving body plays a central role, one can well understand how central the reflection on theatrical costume is. Evidence of this is an ancient tradition of writings and theorizations. Just think of a seminal document in the development of theatrical dance, such as *Lettres sur la danse et sur les Ballets* (1760) by Jean Georges Noverre (1727-1810), and how with it the discourse on costumes practically goes hand in hand with that on dance styles. According to the *Lettres* costumes, indeed, should not only be aesthetically pleasing but also functional, allowing dancers to express their characters freely through movement [10]. Thus, each definition of the characteristics of the identified dance styles corresponds to a precise type of body and costume: the *sérieux* dancer is characterized by a statuesque and slender beauty and broad and elegant movements. A 'gracefulness' in costumes is represented by strong shapes, bows, and high

 ISBN: 978-1-68558-154-1

plumes (Figure 1, *left*). The*demi-caractère* dancer is more dynamic and fast in movements, has clothing that allows greater mobility, less geometrically structured drapes, and brighter colors to emphasize the character (Figure 1, *right*) [11].

Strongly linked to bodies and their identity, dance offers a rich array of evidence on the ability of costumes to influence and condition movement. Just think, as an example, of the pioneering contribution of dancer Isadora Duncan (1877-1927) in opening up to a free mode of movement's expression and the role that soft and shoeless clothing played in these innovations (Figure 2) [12].

The concept of functionality thus arises, in the study of dance costumes, as characterized by a multifaceted dimension linked to poetics, execution techniques, and the aesthetics sought by both the performer and the choreographer. This point explains further the ability, which we previously defined as 'agency,' of the costume to 'tell' the specificities of the performer, the historical period in which he or she lives, together with his or her tastes and trends.

Considering a source like this in the analysis of Dance Heritage is indispensable. Doing so from a perspective that encompasses movement, volumes, and bodies, while also addressing the preservation and valorization of resources, presents a challenge to which the use of technology can offer intriguing insights.

Figure 1: From left to right, Maquette de Costume pour serieux (*Berger*); Maquette de Costume pour Demi-caractaire, by Louis René Boquet, ©Bibliothèque nationale de France.

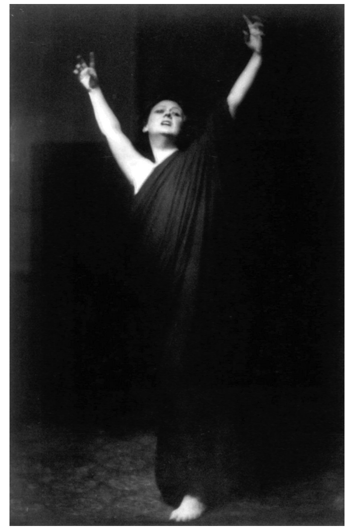

Figure 2: Isadora Duncan performing barefoot during her 1915–1918 American tour, ph. Arnold Genthe, ©Library of Congress.

II. Case Study: Rudolf Nureyev (1938-1993)

In this section, we detail our case study related to the public image of Rudolf Nureyev, examining the data correlated to him and the challenges raised by them.

A. Public Image

Rudolf Nureyev was a famous dancer and choreographer, whose performances, during the 1960s and 1980s, electrified audiences worldwide. Renowned for his technical skill and emotional depth, he rose to fame as a principal dancer with the Kirov Ballet in Leningrad. In 1961, after his defection to the West, he captured global attention and became a global phenomenon [13]. The widespread development of mass media between of the time accompanied the dissemination of his fame, characterized by strong attention not only to dance skills and on-stage presence but also to his public image as a member of the international jet set. Nureyev's media image is marked by this polarization: on one hand, the dance star with exceptional virtuosity and great abilities, and on the other, the style icon, his private life, attire, and gossip. These two aspects influenced his placement within the public imagination.

B. Costumes and Fashion

Recognizing the significant power of his body in conveying artistry, Nureyev meticulously paid attention to costume selection throughout his career. Moments marked as "artist's whims," such as delaying the start of the final act of *Don Quixote* at the Kirov Theatre (because the trousers in the costume did not flatter his leg line), or systematically covering the streets of Paris upon his arrival in 1961 (in search of wigs and costumes) [14], actually reveal much about his artistic and aesthetic languages.

A comparative view of the costumes chosen for his choreographies, for example, shows recurring traits, such as a very short bodice, light-colored shoes, and tights to elongate the figure, wide openings at chest level to emphasize the face and bust proportions, rich decorations along the lines of force, and meticulous control over the final execution. The vision of the original sketches, in this regard, tells us much about the possibilities of intervention that he always reserved for himself [15]. To this it must be added that numerous newspaper articles present him not only as a skilled interpreter but also as a style

 ISBN: 978-1-68558-154-1

icon, frequently pairing these two aspects as shown in Figure 3.

The awareness and attention given to these aspects made Nureyev's public image highly recognizable and mediatized, capable of withstanding the test of time and influencing generations of dancers who adopted many of the aesthetic innovations he introduced. The same discourse naturally extends and reverberates within the broader public sphere and the general public. Just consider the work done to revive the ”Nureyev style” by fashion houses, such as Etro [16] or Dior for their men's collections [17].

Figure 3: Rudolf Nureyev on the cover of ”Observer” in 1976 (Paris, ”Noureev Collection”/Press/1975-79, ©Centre national de la danse)

C. Challenges

Due to its comprehensive exploration of costume agency, Nureyev serves as an ideal case study. In narrating the dancer's legacy, the costume indeed plays a central role, albeit posing some difficulties to both the scholar and the user. These, to offer a synthesized view, can be attributed to two strongly interconnected categories:

- Preservation
- Valorization

The international nature of Nureyev's career has led to a strong dislocation of sources, scattered among museum institutions, theatrical archives, and private collections worldwide. The value of these objects, along with the fragility associated with costumes already heavily worn from intense use, requires the utmost care in preservation operations. In both cases, the accessibility of the heritage can only be affected by inevitable limitations, with repercussions in terms of valorization.

Masterful in this regard is the work of the Fondation Rudolf Noureev (RNF) and the Centre National du Costume de Scène in Moulins (CNCS), which have dedicated an entire exhibition section to the ”Collection Noureev” [15] consisting of costumes, scene sketches, everyday clothing, personal items, and furniture from his Parisian apartment. The collection provides a tangible example of the full range of documentary resources necessary to activate costume agency: photographs, sketches, videos, but also newspaper articles, and texts revolving around the performances (programs, posters, etc.). Hence, in order to explore potential examples that ensure both preservation and valorization, we will review and apply in real cases those paradigms allowing the reconstruction of 3D objects from 2D images. The chosen primary sources are among the most widespread and ready-to-use for what concerns Nureyev: photographs featured in international print newspapers.

III. From 2D images to 3D objects

A. Related works

Transitioning from 2D images to 3D models involves various methods, which are listed below.

Among the most widespread approaches, we mention Structure-From-Motion (SfM) [18], which reconstructs 3D geometry from a series of 2D images by analyzing their relative positions and orientations. In particular, SfM methods first detect and match key points between pairs of images. The camera poses are estimated and 3D coordinates are triangulated by intersecting the rays back-projected from the cameras. Finally, they provide a dense 3D model with texture mapping. SfM is advantageous for its simplicity and ability to handle large datasets, however, it struggles with accuracy in complex scenes and requires considerable computational resources. Furthermore, multi-view stereo (MVS) [19] methods have been developed. Different from SfM, this class of methods estimates depth information by analyzing multiple images of the same scene. MVS offers improved accuracy compared to SfM, particularly in detailed environments, but it can be computationally intensive and sensitive to image quality and viewpoint variations. Photogrammetry is another widely used method for creating 3D models from 2D images. The process involves analyzing the geometric properties of images taken from different viewpoints to reconstruct the three-dimensional structure of objects or scenes. Photogrammetry (Ph) [20] offers advantages, such as scalability, as it can be applied to images captured from consumer-grade cameras or drones, and it provides detailed surface information. However, photogrammetry also has limitations, as it can be sensitive to factors like lighting conditions, camera calibration, and image quality, which may affect the accuracy of the resulting models. Additionally, it requires careful planning and processing to ensure reliable results, and capturing images from multiple viewpoints can be time-consuming. Gaussian splatting (GS) [21] is a technique commonly used in computer graphics and computer vision for rendering 3D scenes from 3D point cloud data. In Gaussian splatting, each point in the point cloud is represented as a Gaussian distribution centered at its position in 3D space. When rendering the scene, these Gaussian distributions are ”splat” onto a 2D image plane, where they contribute to the pixel intensity based on their spatial extent and distance from the pixel. This process effectively blends the contributions of nearby points in the point cloud, resulting in smooth and continuous renderings. Gaussian splatting is particularly useful for visualizing dense point cloud data captured from sources like LiDAR sensors or depth cameras. So far, AI-based approaches have been used, for example, to assist SfM, MVS, GS or Photogrammetry approaches in solving tasks like feature extraction, matching, and reconstruction, leading to more accurate and efficient 3D model generation, which by leveraging AI can mitigate some of their traditional limitations and offer improved results across a range of applications. However, AI methods have been developed also to directly estimate 3D meshes. Pixel-aligned Implicit Function (PIFu) [22], employs a neural network architecture to predict a voxelized representation of the 3D object directly from the 2D input image. This method stands out for its ability to generate highly detailed and accurate 3D reconstructions even with one single image, capturing fine-grained surface details through pixel-aligned implicit functions. Moreover, Neural Radiance Fields (NERF) [23] are capable of representing scenes as continuous volumetric functions describing radiance at any given 3D point. By training a neural network to predict this function from 2D images and corresponding camera poses, NeRF achieves photorealistic reconstructions with high fidelity across viewpoints.

B. Application of Gaussian Splitting models

In Figure 4, we report three views of the reconstructed 3D model obtained by applying a Gaussian splitting model to the 2D image of the Double for the role of Basilio in *Don Quixote*. Despite being generated from a single 2D view, the 3D reconstruction exhibits sufficient accuracy and captures the main details. Furthermore, by restoring its three-dimensionality, the technology employed allowed for a 360° view of the object (rebuilt thanks to AI), thus enabling it to perform under conditions as close as possible to the original, intended for continuous motion viewing.

In addition, we crop from newspaper pages a hat and a doublet used by Nureyev and we provide 3D reconstructions from one image using a Gaussian splitting model. In Figure 5 and 6, we report the 2D crop and two views of the 3D objects, one related to the repertoire of stage costumes, the other to a fashion item.

Working from newspapers, in the case of Nureyev, is crucial. As we have mentioned, his figure has been heavily mediatized, making him highly present in this type of documentation. However, this choice also presents significant benefits for adding to a costume heritage that would otherwise be scarcely accessible because stored and sold at auctions or dispersed. In this way, it is possible to expand existing collections and, in cases where costumes are already allocated, contribute to describing the resource by adding dates of use, and performances, and tracing their media coverage.

IV. Conclusion

This contribution has highlighted the pivotal role of costumes in dance heritage, exemplified through the case study of Rudolf Nureyev. Costumes serve as dynamic agents, conveying narratives, aesthetic preferences, and historical contexts, while also influencing movement and performance. Nureyev's career exemplifies how meticulous attention to aspects related to costume contributed to the construction of the imagery associated with him and thus to the perpetuation of his artistic legacy. In Nureyev's case, this extends to the broader realm of clothing and fashion choices, which have greatly shaped his public image. However, the preservation and valorization of those items pose significant challenges: How is it possible to integrate the volumetric and dynamic dimensions of costumes designed for movement? How can their integrity be preserved without sacrificing communication with the audience? How can widely available sources be transformed into fonts of new information? Innovative approaches such as 3D reconstruction techniques offer a valid example to address these questions. By employing photogrammetry and Gaussian splatting models, this study demonstrates the potential of technology in enhancing preservation efforts and enabling closer examination of costume design and detail. Once reconstructed in 3D, these objects offer significant usability, overcoming the fragility associated with their uniqueness and value, enhancing comparative viewing of heritage assets, detailed examination, usability, and preservation. Appropriately processed digital objects also contribute to this effort. The interdisciplinary approach used in this study further sheds light on the intricate interplay between dance, costume, and technology, underscoring the importance of promoting dance heritage for future generations and doing so by collaborating with experts from different domains. Considering future investigations, we will adopt different neural rendering paradigms, including also neural radiance fields, and compare quantitatively and qualitatively the approximated 3D models, involving dance experts.

Figure 4: Views of the 3D model of the Doublet for the role of Basilio in *Don Quixote*, 1979. Costume by Nicholas Georgiadis. ©CNCS / Photo Pascal Francois.

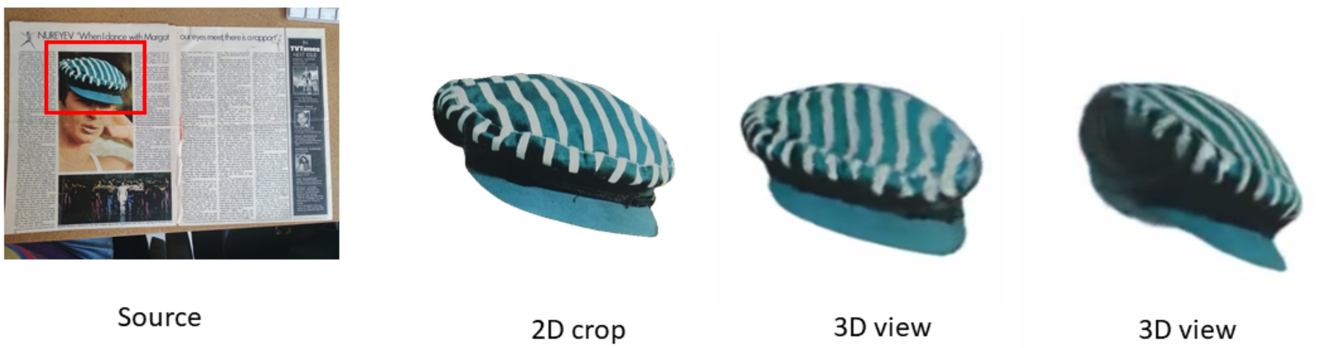

Figure 5: Views of the 3D model of a hat taken from a magazine (Paris, ”Noureev Collection”/Press/1972, ©Centre national de la danse).

Figure 6: Views of the 3D model of Nureyev's *Swan Lake* Doublet for ”The Muppet's Show”, starting from a photo on a magazine (Paris, ”Noureev Collection”/Press/1972, ©Centre national de la danse).

REFERENCES

[1] J. Adshead-Lansdale and J. Layson, *Dance history: An introduction*. Routledge, 2006.

[2] E. Giannasca, “Dance in the ontological perspective of a document theory of art,” *Danza e ricerca. Laboratorio di studi, scritture visioni*, vol. 10, no. 10, pp. 325–346, 2018.

[3] H. Green, “Cultural history and the material (s) turn,” *Cultural History*, vol. 1, no. 1, pp. 61–82, 2012.

[4] T. J. Schlereth, *Cultural history and material culture: Everyday life, landscapes, museums*. University Press of Virginia, 1992.

[5] A. S. Martin and J. R. Garrison, “Shaping the field: The multidisciplinary perspectives of material culture,” *American material culture: The shape of the field*, pp. 1–20, 1997.

[6] V. Fuchsberger, M. Murer, and M. Tscheligi, “Materials, materiality, and media,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 2853–2862, 2013.

[7] I. Woodward, “Understanding material culture,” *Understanding Material*

Culture, pp. 1–200, 2007.
[8] C. Tilley, "Materiality in materials," *Archaeological dialogues*, vol. 14, no. 1, pp. 16–20, 2007.
[9] D. Barbieri, *Costume in performance: materiality, culture, and the body*. Bloomsbury Publishing, 2017.
[10] P. Dotlačilová, *Costume in the Time of Reforms: Louis-René Boquet Designing Eighteenth-Century Ballet and Opera*. PhD thesis, Stiftelsen för utgivning av teatervetenskapliga studier (STUTS), 2020.
[11] F. Pappacena, *La danza classica: le origini*. Gius. Laterza & Figli Spa, 2014.
[12] E. Cervellati *et al.*, *Storia della danza*. Pearson Italia spa, 2020.
[13] R. Nureyev, "Nureyev: an autobiography with pictures," *(No Title)*, 1962.
[14] J. Kavanagh, *Rudolf Nureyev: the life*. Penguin UK, 2013.
[15] M. Kahane, *Rudolf Noureev 1938-1993. Costumes et photographies.* CNCS Editions du Me'ce'ene, 2009.
[16] L. Leithch, "Etro Fall 2021 Ready to-wear (Published 2021) — vogue.com." https://www.vogue.com/fashion-shows/fall-2021-ready-to-wear/etro, 2021.
[17] S. Mower, "Dior Men Fall 2024 Menswear (Published 2024) — vogue.com." https://www.vogue.com/fashion-shows/fall-2024-menswear/dior-homme, 2024.
[18] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds, "'structure-from-motion'photogrammetry: A low-cost, effective tool for geoscience applications," *Geomorphology*, vol. 179, pp. 300–314, 2012.
[19] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*, vol. 1, pp. 519–528, IEEE, 2006.
[20] T. Schenk, "Introduction to photogrammetry," *The Ohio State University, Columbus*, vol. 106, no. 1, 2005.
[21] G. Chen and W. Wang, "A survey on 3d gaussian splatting," *arXiv preprint arXiv:2401.03890*, 2024.
[22] S. Saito, Z. Huang, R. Natsume, S. Morishima, A. Kanazawa, and H. Li, "Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 2304–2314, 2019.
[23] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

Photogrammetry and 360° Virtual Tours: Differences, Relevance, and Future Possibilities

Irene Calvi
Department of the Arts
University of Bologna
Bologna, Italy
email: irene.calvi2@unibo.it

Eleonora Stacchiotti
Department of the Arts
University of Bologna
Bologna, Italy
email: eleonora.stacchiotti2@unibo.it

Pasquale Cascarano
Department of the Arts
University of Bologna
Bologna, Italy
email: pasquale.cascarano2@unibo.it

***Abstract*—This paper examines the impact of technology integration in fashion retail, specifically focusing on virtual tours using 360° cameras and photogrammetry. Through case studies of Ralph Lauren and Dolce & Gabbana, it compares the strengths and limitations of each approach, emphasizing user experience and brand representation. While 360° tours offer accessibility, photogrammetry provides immersive experiences and detailed visualizations. The paper addresses challenges like content relevance and technological obsolescence, proposing strategies for innovation. It also discusses a project at the University of Bologna's VARLab for AEFFE s.p.a., exploring the future integration of virtual tours into transmedia storytelling. Additionally, the study investigates the role of VR in luxury fashion retail, emphasizing emotional impact and aesthetic engagement, presenting a preliminary report on immersive VR video design.**

***Index Terms*—Virtual tours; Fashion retail; Photogrammetry; Brand heritage; Immersive experiences.**

I. Introduction

In recent years, there is an abundance of studies that have explored the evolution of technology in fashion, dedicating particular attention to the retail context. The shopping experience has been defined as made by multiform sensory and informational solicitations [1] that aim to engage customers and enhance their overall satisfaction. However, confusion within the marketing and communications team has led to a proliferation of creations by brands using similar keywords without much differentiation [2], [3]. This has resulted in an abundance of "virtual tours", "showrooms", and "museums". Instead of providing an overview of these experiences, let's focus on understanding the differences between two types of virtual tours: those realized through a 360° camera and those realized through photogrammetry. For many years, 360° videos were more commonly used due to their ease of production. On the other hand, photogrammetry often yields better results in terms of accessibility and immersiveness [4]. These technologies are considered examples of Desktop Virtual Reality known for enhanced 3D visualization capabilities compared to Immersive Virtual Reality. The terminology surrounding virtual tours can be confusing; here it refers to a tour within a virtual reproduction representing an actual physical space with depth perception [5]–[7]. To provide further clarification on this choice – only two case studies will be examined corresponding with either 360° cameras or photogrammetry techniques allowing for greater insights into their respective strengths and limitations. We then focused on 360° cameras and video, and in general, Virtual Reality (VR) and Augmented Reality (AR) paradigms in fashion. Fashion functions as an interpreter of societal change, adeptly deciphering subtle cues and trends, and subsequently reintroducing them into the commercial forefront through a process of remixing and refinement. Nowadays, this perpetual cycle of interpretation and adaption is performed by adopting modern technologies [8].

VR has emerged as a pivotal tool in enhancing the online retail experience, bridging the gap between physical and digital realms [9]. The Metaverse [10] is an interconnected virtual universe that can transform the way individuals interact and transact within digital spaces, with VR serving as a gateway to this new frontier [11]. Defined as "the use of computer simulation that enables interaction with a virtual, three-dimensional, visual environment through digital representation" [12], VR transports users into digitally reconstructed worlds via Head-Mounted Displays (HMDs), creating a sense of presence and immersion [13]. Projections indicate a steep rise in VR headset adoption, with the majority of internet users projected to incorporate VR into their daily routines within the next decade [14]. Within the realm of retail, VR applications are rapidly evolving, offering novel avenues for consumer engagement and experiential marketing [15], [16]. By integrating exclusive content and VIP experiences inaccessible in the physical realm, VR further enhances the allure of luxury fashion brands [17]. It is within this dynamic landscape that the present study unfolds—a quest to explore the potential of immersive 360° VR videos in eliciting aesthetic experiences among users. As part of a broader research endeavor aimed at investigating and developing Extended Reality (XR) prototypes tailored for luxury fashion brands, this paper serves as a preliminary report detailing the design process of an immersive VR experience. By documenting the initial steps undertaken, the paper sets the stage for further exploration into the possibilities afforded by XR technologies within the realm of fashion retail with a focus on the emotional impact that XR experiences can provoke in the users. This manuscript will conclude with an ongoing project carried out within the Virtual and Augmented Reality

Laboratory (VARLab) [18] of the University of Bologna for the luxury company AEFFE s.p.a. [19].

II. Case Studies

In the following section, we present two case studies of prominent fashion brands, Ralph Lauren and Dolce & Gabbana, which leverage virtual experiences to engage customers and showcase their products.

A. *Case Study 1: Ralph Lauren Virtual Experiences*

The first brand to be mentioned is Ralph Lauren, a renowned fashion house founded in 1967. The brand is famous for its distinctive polo shirts, tailored suits, and an iconic logo featuring a polo player on horseback. With a focus on timeless designs, Ralph Lauren caters to a discerning clientele seeking refined and sophisticated attire. The brand's aesthetic seamlessly blends traditional Americana with high-end craftsmanship, creating an aspirational lifestyle and sense of prestige. Offering an extensive range of products including clothing, accessories, fragrances, and home furnishings, Ralph Lauren continues to shape the landscape of contemporary fashion and luxury lifestyle. Their webpage dedicated to virtual stores is not connected to their e-commerce shop and changes based on the audience. Generally, the virtual tours are defined "RL Virtual Experience" or "Virtual Flagship Store", see Figure 1 and Figure 2. Three distinct platforms that cater to different geographical markets were considered: the first targets a global audience, the second caters primarily to a US audience the third is directed towards the UK market, featuring stores in Milan and London. Each URL shares a common feature, the 888 House, providing an immersive and experimental space to display the collection, while the global URL emphasizes characteristics specific to the winter holidays, including in-store decorations, a gifts section, and Christmas background music.

Overall, such platforms exhibit characteristics tailored to enhance user engagement and navigational ease, integrating an audio component for realism. Utilizing a drag-around feature coupled with anchored arrows, users are afforded seamless movement throughout the virtual space, facilitating exploration of various sections within the store environment. Complementing this navigational aid, a map sign positioned on the right side of the interface serves as a helpful guide, elucidating the spatial layout and enabling users to discern different sections efficiently. Interactive elements were strategically positioned throughout the virtual environment, offering users the opportunity to delve deeper into showcased products [20], [21]. If these objects remain accessible online, they seamlessly redirect users to their corresponding pages for online shopping, complete with detailed product descriptions and prices [22]. In instances where online access is no longer available, users are seamlessly redirected to the main page, ensuring continuity of the browsing experience. Notably, within the preview of each object, pertinent details including price and descriptive information are prominently displayed, facilitating informed decision-making and enhancing user engagement with the

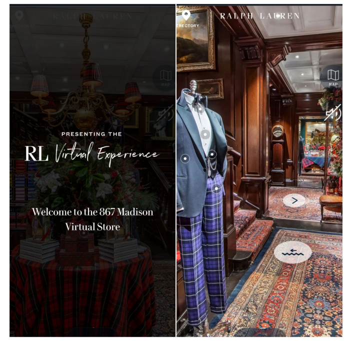

Fig. 1: RL virtual experience homepage and interiors of 867 Madison store.

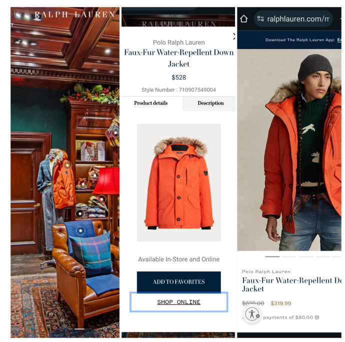

Fig. 2: Stills of RL 867 Madison Virtual Store showing a jacket on the mannequin, the information showing when clicked on it and in the e-commerce from left to right.

virtual store environment. In conclusion, these meticulously crafted virtual tours embody a multifaceted approach aimed at providing users with an immersive and informative shopping experience within the realm of Ralph Lauren's storied brand universe.

B. *Case Study 2: Dolce & Gabbana Virtual Boutiques*

The second case study is devoted to Dolce & Gabbana's choice of recreating their stores with photogrammetry. The Italian luxury fashion house was established in 1985 and embodies opulence and sensuality in its creations. Renowned for its bold and glamorous designs, the brand exudes a distinct Mediterranean flair infused with Italian heritage and craftsmanship. Dolce & Gabbana's virtual boutiques stand out as a hallmark of innovative online retailing, offering patrons a distinct and immersive experience via a dedicated website distinct from their e-commerce platform. Upon navigating the main menu, users are greeted with several options, from previews of upcoming virtual boutique unveilings to the convenience of booking appointments, accessing the Women's and Men's fashion shows for FW2022, utilizing the store locator, and revisiting past virtual boutiques. This designated section, defined as "Relive Virtual Boutiques," serves as a gateway to explore eleven distinct venues spanning the globe, including iconic locations, such as Venezia, Roma, Osaka, Chadstone, Shanghai, Paris, Miami, Las Vegas, Cannes, Seoul,

Fig. 3: Dolce & Gabbana virtual boutiques of Venice (upper half) and Rome (bottone half) venues.

and Tokyo. Each venue offers a bespoke lens through which to perceive the brand's rich tapestry of identity and aesthetics, from the grandeur of historical European settings to the pulsating energy of urban metropolises, see Figure 3. While some of the more recent virtual venues may feature embedded links directing users to the e-commerce platform, it's noted that the functionality of these links may vary, with some potentially rendered inactive due to expiration. The presence or absence of such links in previous iterations remains a matter of uncertainty, hinting at the dynamic evolution of the virtual boutique landscape over time.

A cornerstone of the virtual boutique experience lies in its utilization of photogrammetry, enabling the meticulous mapping of spatial dimensions and textures that afford users a degree of freedom in navigation and exploration. This platform immerses users in a sensory journey through the heart of Dolce & Gabbana's design ethos. While certain venues like Venezia and Roma serve as showcases of the brand's Italian heritage, it's noteworthy that the most contemporary settings bear subtle yet unmistakable traces of this identity. Whether in the ornate flourishes adorning architectural details or the evocative styling of merchandise displays, these recurring elements serve to uphold a sense of continuity and coherence across the diverse array of virtual boutiques, ensuring a seamless brand experience for patrons worldwide [23].

III. Case Studies Analysis

An in-depth examination of the positive and negative attributes inherent in the utilization of both a 360° virtual tour and a photogrammetry virtual tour within the context of fashion showrooms illuminates various nuanced aspects. Despite their differences, these modalities share a common functionality in offering a dedicated section for booking appointments, thereby facilitating seamless customer interaction and engagement. Delving into the advantageous aspects, these virtual tours emerge as potent tools in shaping the multifaceted image of a brand, serving as dynamic contributors to its identity construction. By providing immersive experiences that transcend the confines of traditional retail environments, they afford consumers a novel and explorative mode of shopping, transcending geographical barriers and time constraints. Moreover, these tours offer a unique opportunity for customers to delve into the architectural intricacies of each showroom, with some establishments boasting more distinctive and culturally resonant designs than others. This architectural exploration not only enhances the overall shopping experience but also fosters a deeper appreciation for the brand's aesthetic diversity and heritage.

Despite notable advantages, virtual tours face challenges such as presenting static images of the past and requiring constant updates to remain relevant amid evolving consumer preferences. The rapid pace of technological advancement and ongoing experimentation render these tours susceptible to obsolescence, necessitating adaptation to maintain efficacy. Given their inception during the pandemic, there is a pressing need for virtual tours to evolve and align with brand missions. This entails significant investments in time and resources to ensure continued relevance in facilitating brand objectives. Addressing these challenges necessitates a proactive approach towards innovation and adaptation, wherein embracing emerging technologies and incorporating interactive features can enhance the immersive nature of these virtual tours, elevating the overall customer experience. By integrating dynamic content and personalized recommendations, fashion brands can ensure ongoing engagement and relevance, catering to the diverse preferences and tastes of modern consumers. Moreover, forging strategic partnerships with tech companies and leveraging data analytics can provide valuable insights into consumer behavior and preferences, guiding informed decision-making and content development efforts.

In this context, Photogrammetry and 360°-video technology stand as two distinct approaches for reconstructing a virtual 3D environment to create immersive experiences. While they consider different paradigms to achieve this goal, the strategic implementation of these technologies can significantly augment brand visibility and customer engagement in the competitive fashion retail landscape.

Photogrammetry utilizes 2D photographs to create detailed 3D models through specialized software like Agisoft Metashape, RealityCapture, or Pix4D, requiring precise photographic capture and good lighting. This method is ideal for sectors like virtual archaeology and civil engineering but can be complex and time-consuming. 360° video technology records panoramic videos for interactive viewing, offering a smoother visualization and cinematic experience suitable for virtual tours. Unlike photogrammetry, it requires less time and technical skills but lacks detailed 3D models. Both approaches have advantages in showcasing fashion showrooms and enhancing brand visibility, yet face challenges. By acknowledging limitations and embracing innovation, fashion brands can maximize the potential of virtual platforms in an evolving digital landscape.

IV. Adopting 360° videos to assess emotional drivers

Considering our previous considerations for 360° videos, we here proceed by crafting an exclusive and immersive

 ISBN: 978-1-68558-154-1

experience, aimed at democratizing access to rare and coveted items within the fashion realm, to cultivate a sense of personal connection between consumers and the products they admire, while simultaneously fostering a vibrant community around the brand. In contemporary fashion marketing, the useof brand ambassadors has become a prevalent strategy for launching and promoting new collections [24]. This phenomenon is particularly significant in today's digital age, where social media platforms serve as central hubs for fashion discourse and engagement [25]. To explore the potential of this approach, a low-cost prototype was conceived to facilitate initial experiments, with 360° videos. The prototype aimed to create a one-to-one immersive encounter wherein a dancer or performer, outfitted in a rare and iconic garment from a luxury brand, moves gracefully around a stationary camera which acts as the user's point of view. The prototype intentionally minimized direct interaction and verbal communication between the performer and the observer to focus on fostering a profound connection between the viewer and the attire itself. The captured 360° video was intentionally short and will be presented to a test group, allowing for an assessment of its practicality and its impact on users. Furthermore, in a bid to explore the dynamic interplay between reality and virtual reality, the prototype will be tested in the same physical environment where the 360° video was filmed. This approach will provide valuable insights into how users perceive and engage with the virtual representation of a physical space, further informing the development of future iterations of the project. By combining elements of experiential marketing, immersive technology, and brand storytelling, the project aims to pave the way for innovative approaches to consumer engagement within the fashion industry. Through experimentation and iteration, it seeks to uncover new opportunities for brands to connect with consumers on a deeper level.

V. METHOD

In the following section, we inspect the intricacies of measuring aesthetic emotions elicited by immersive 360° video experiences, detailing the process, methodologies, and expected outcomes of the research endeavor.

A. Making of the 360° video

The process of capturing 360° video with Insta360 cameras involves recording the video, transferring the files, and stitching, which is the process of combining different camera perspectives to create a seamless panoramic video. The video lasts 10 minutes based on the average attention span of users [26], [27], [28]. Here are the details of the process:

1) Recording and transferring the video: using an Insta360 camera to record the 360-degree video capturing images/videos from all directions simultaneously, recording the entire surrounding environment.
2) Stitching: after transferring the files to the processing device, the stitching process is required. The stitching process combines the different camera perspectives in the 360-degree camera to create a seamless panoramic video. Insta360 stitching software, whether integrated into the cameras or provided separately, analyzes images from different lenses and blends them to ensure a smooth transition between different views.
3) Editing: after stitching, you can make any necessary edits to the video using video editing software. This phase allows you to add effects, adjust colors and contrasts, add audio, and make other customizations to 360-video. In the case of our video, the tripod on which the Insta360 camera was always present in the frames. Therefore, a mask was applied to hide the tripod to reach a sense of realism to the experience.
4) Exporting the final video: once the stitching and the optional editing are complete, the final video is exported and uploaded into a structure for environmental images to be visualized with the headset.

In Figure 4, we report some screenshots taken from the 360° video produced in the described process.

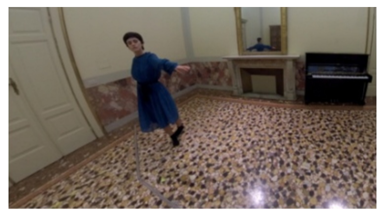

Fig. 4: Screenshot taken from the produced 360° video.

B. Selection of target audience

Once the video production is complete, it will undergo evaluation by users spanning diverse age demographics and varying levels of digital literacy. The primary objective is to discern any potential influence of user age and familiarity with immersive technologies on their perception of the experience. Although the current trend of leveraging new technologies to appeal to younger consumer segments, it is imperative to ensure accessibility and usability for older generations as well. This is particularly significant considering the important presence of millennials and older individuals within the luxury market [29], [30], [31], [32]. The final sample size for user testing is estimated to comprise approximately 10 to 30 individuals, representing a broad spectrum of age groups. This number has been determined as a balance between the need to gather sufficient data from a diverse sample and the practical constraints of the evaluation phase. It is anticipated that this sample size will exceed 10, as previous research has indicated that a minimum of 10 testers is sufficient for identifying meaningful correlations across various parameters in interface design [33], [34], [35].

C. Measuring Aesthetic Emotions

The first objective is to evaluate the aesthetic experience of the user. Cognitive scientists have tried to implement direct and indirect measures for aesthetic emotional experience in experiments in empiric aesthetics. Given that making appropriate aesthetic judgments is not fixed but is modulated by daily behavioral habits [36], Mastandrea et al. reflected on the account of the relationship between aesthetic emotion and physical-psychological well-being [37]. The behavioral-qualitative investigation of aesthetic experiences focuses on the measurement of self-reported emotions. Others suggested an analysis of existing measures for aesthetic emotions within specific domains of artistic productions [20], [21], [38]. They developed the Aesthetic Emotions Scale (AESTHEMOS) structured around groups of emotions:

- **Prototypical aesthetic emotions like the feeling of beauty**: Being moved; Awe; Fascination;
- **Epistemic and emotions like interest**: Joy; Relaxation; Vitality; Negativity;

Hereby, I report a description of the survey that will be given to the sample users.

a) Aesthetic Emotions Scale (AESTHEMOS): The AESTHEMOS can be used to assess either the intensity of aesthetic emotions (e.g., for studying momentary aesthetic experience or the experience of a specific stimulus, such as a picture, poem, piece of music, or film scene) or the frequency of experiencing aesthetic emotions during a more prolonged aesthetic experience (e.g., for studying an event as a whole, such as an entire art exhibition, theater performance, or a walk through-nature).

In this proposal, the scale adopted matches exactly the AESTHEMOS introduced by Schidler et al. [38], assessing the intensity of all the emotional aforementioned constructs. In particular, each emotional construct is detailed in pairs of questions.

For the frequency version, the following modifications need to be made: (1) Rating instruction: How often did you feel this emotion? (2) Rating scale: A 5-point Likert Scale where 1 stands for never and 5 very often; (3) Instructions: Which emotional effect did the experience have on you? For each emotion listed below, please mark the response category that best matches your personal experience. Please only indicate how you felt. Do not characterize the emotions expressed in the experience if you did not feel them yourself.

D. Expected results and directions for further research

We expect to organize the testing sessions as follows. First of all, the users will be asked to fill a welcome form to declare their demographic data and their initial emotional state with a short Self-Assessment Manikin [39]. After that, they will be shown the immersive 360° video on a VR headset and they will be asked to answer the survey reported above in Table 1. Based on the questions of the AESTHEMOS scale, this study forecasts to have a starting impression of the emotional impact of an immersive 360° video with low human interaction on the viewer. Among other aims, we hope to understand whether there is a correlation between demographic data, digital literacy and the intensity of experienced emotions. This will lead to the development of further research, that can be implemented in several ways. For example, we could divide the sample users according to different age groups to better identify the correlation between age and VR experience. Or, we could develop other experiments to determine what sharpens the sensations felt during the test - i.e., the design of a digital twin or its motion, the interaction between user and digital twin, the computer-generated environment vs. the location of the experiment. However, considering that the survey must be filled moments after the experience, the result could be influenced by many external factors. For this reason, in the longer term, it is advisable to start experimenting with methodologies which - among others - evaluate emotional and affective reactions to stimuli by measuring, collecting and elaborating biometric data such as heartbeat, sweating, blood pressure or by finding ways to track the movement of the eyes - all elements that need to be monitored during the experience.

VI. Conclusion and further possibilities

As part of the PNRR project of Spoke 1, the authors and other colleagues have recently made a 360° virtual tour of the Moschino showroom in Via della Spiga in Milan. Moschino is one of AEFFE s.p.a.'s four brands, and the store had a well-designed and planned reopening in 2022 followed by another store opening in Rome. The new challenge embraced is to prototype a space that is less focused on being just a vitrine for e-commerce and more as part of an immersive narrative experience. To achieve this goal, two elements will be leading the experience: the store design and the window display. The newly designed stores, initially envisioned by former creative director Jeremy Scott, may undergo further evolution under potential new creative direction from Adrian Appiolaza, featuring elements reminiscent of Franco Moschino's bold and playful designs while paying homage to his studio through unique displays of precious accessories. Virtual tours capture specific moments featuring distinctive vitrine props from each collection over a time frame creating an engaging exploration connecting future retail window displays with implications extending into fashion studies through showcasing brand heritage. The ongoing objective is to integrate virtual tours into transmedia storytelling for enhanced visual retail communication and increased customer engagement with brand heritage.

To achieve such a goal, we will preliminarily gather comprehensive data regarding users' demographic profiles and validate their emotional states, and their responses to immersive 360° video experiences, aiming to elucidate the emotional impact of the VR environment through the utilization of the AESTHEMOS scale. This would set the stage for future research aimed at exploring factors influencing users' emotional responses, including age-related nuances, digital twin design, interaction dynamics, and environmental factors. Considering the latter, future research directions may incorporate methodologies such as biometric data capture to provide additional insights into users' emotional and affective reactions.

In conclusion, while the initial testing sessions offer valuable insights into users' immediate emotional experiences, future research endeavors will delve deeper into the intricate dynamics of emotional engagement within immersive environments, in particular, applied to the realm of luxury fashion retail.

ACKNOWLEDGMENT

This study was carried out within the MICS (Made in Italy – Circular and Sustainable) Extended Partnership and received funding from the European Union Next-GenerationEU (Piano Nazionale di ripresa e resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.3 - D.D. 1551.11-10-2022, PE00000004).

REFERENCES

[1] V. M. Iannilli and V. Linfante, "Exploring disrupting scenarios in the fashion retail and communication paradigms," *Luxury Studies: The In Pursuit of Luxury Journal*, vol. 1, no. 1, pp. 45–65, 2022.

[2] M. Herz and P. A. Rauschnabel, "Understanding the diffusion of virtual reality glasses: The role of media, fashion and technology," *Technological Forecasting and Social Change*, vol. 138, pp. 228–242, 2019.

[3] B. Jin, G. Kim, M. Moore, and L. Rothenberg, "Consumer store experience through virtual reality: its effect on emotional states and perceived store attractiveness," *Fashion and Textiles*, vol. 8, pp. 1–21, 2021.

[4] R. Boardman, C. E. Henninger, and A. Zhu, "Augmented reality and virtual reality: new drivers for fashion retail?," *Technology-Driven Sustainability: Innovation in the Fashion Supply Chain*, pp. 155–172, 2020.

[5] A. M. Rodríguez-López, S. Rubio-Valdehita, and E. M. Díaz-Ramiro, "Influence of the covid-19 pandemic on mental workload and burnout of fashion retailing workers in spain," *International Journal of Environmental Research and Public Health*, vol. 18, no. 3, p. 983, 2021.

[6] Y. Suryani, M. A. Ayu, and M. J. Wahyuni, "Sense of presence in a virtual reality fashion show 360 video," in *2020 6th International Conference on Computing Engineering and Design (ICCED)*, pp. 1–6, IEEE, 2020.

[7] A. S. Sina and J. Wu, "The effects of retail environmental design elements in virtual reality (vr) fashion stores," *The International Review of Retail, Distribution and Consumer Research*, vol. 33, no. 1, pp. 1–22, 2023.

[8] A. for Industries, "How immersive technologies are changing the retail landscape," 2021.

[9] D. Grewal, S. Motyka, and M. Levy, "The evolution and future of retailing and retailing education," *Journal of Marketing Education*, vol. 40, no. 1, pp. 85–93, 2018.

[10] M. Alcañiz, E. Bigné, and J. Guixeres, "Virtual reality in marketing: a framework, review, and research agenda," *Frontiers in psychology*, vol. 10, p. 1530, 2019.

[11] B. Shen, W. Tan, J. Guo, L. Zhao, and P. Qin, "How to promote user purchase in metaverse? a systematic literature review on consumer behavior research and virtual commerce application design," *Applied Sciences*, vol. 11, no. 23, p. 11087, 2021.

[12] F. Biocca, "Will simulation sickness slow down the diffusion of virtual environment technology?," *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 3, pp. 334–343, 1992.

[13] T. B. Sheridan, *Telerobotics, automation, and human supervisory control*. MIT press, 1992.

[14] P. Rosedale, "Virtual reality: The next disruptor: A new kind of worldwide communication," *IEEE Consumer Electronics Magazine*, vol. 6, no. 1, pp. 48–50, 2016.

[15] A. Javornik, "Augmented reality: Research agenda for studying the impact of its media characteristics on consumer behaviour," *Journal of Retailing and Consumer Services*, vol. 30, pp. 252–261, 2016.

[16] H. McCormick, J. Cartwright, P. Perry, L. Barnes, S. Lynch, and G. Ball, "Fashion retailing–past, present and future," *Textile Progress*, vol. 46, no. 3, pp. 227–321, 2014.

[17] S. J. Kim, "Virtual fashion experiences in virtual reality fashion show spaces," *Frontiers in Psychology*, vol. 14, p. 1276856, 2023.

[18] VARLab, "Varlab website." https://site.unibo.it/varlab/en.

[19] AEFFE, "Aeffe website." https://aeffe.com/it/.

[20] L. Stacchio, C. Scorolli, and G. Marfia, "Evaluating human aesthetic and emotional aspects of 3d generated content through extended reality," 2022.

[21] C. Scorolli, E. N. Grasso, L. Stacchio, V. Armandi, G. Matteucci, and G. Marfia, "Would you rather come to a tango concert in theater or in vr? aesthetic emotions & social presence in musical experiences, either live, 2d or 3d," *Computers in Human Behavior*, vol. 149, p. 107910, 2023.

[22] E. Morotti, L. Stacchio, L. Donatiello, M. Roccetti, J. Tarabelli, and G. Marfia, "Exploiting fashion x-commerce through the empowerment of voice in the fashion virtual reality arena: Integrating voice assistant and virtual reality technologies for fashion communication," *Virtual Reality*, pp. 1–14, 2022.

[23] L. Stacchio, M. Perlino, U. Vagnoni, F. Sasso, C. Scorolli, and G. Marfia, "Who will trust my digital twin? maybe a clerk in a brick and mortar fashion shop," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 814–815, IEEE, 2022.

[24] D. Suleman, Y. Suyoto, R. Sjarief, S. Sabil, S. Marwansyah, P. Adawia, and A. Puspasari, "The effects of brand ambassador and trust on purchase decisions through social media," *International Journal of Data and Network Science*, vol. 7, no. 1, pp. 433–438, 2023.

[25] S.-C. Chu and Y.-K. Seock, "The power of social media in fashion advertising," 2020.

[26] K. L. Nielsen, "Students' video viewing habits during a flipped classroom course in engineering mathematics," 2020.

[27] F. Loh, F. Wamser, F. Poignée, S. Geißler, and T. Hoßfeld, "Youtube dataset on mobile streaming for internet traffic modeling and streaming analysis," *Scientific Data*, vol. 9, no. 1, p. 293, 2022.

[28] S. Yang, D. Brossard, D. A. Scheufele, and M. A. Xenos, "The science of youtube: What factors influence user engagement with online science videos?," *Plos one*, vol. 17, no. 5, p. e0267697, 2022.

[29] C. Amatulli, G. Guido, and R. Nataraajan, "Luxury purchasing among older consumers: exploring inferences about cognitive age, status, and style motivations," *Journal of Business Research*, vol. 68, no. 9, pp. 1945–1952, 2015.

[30] G. Guido, C. Amatulli, and A. M. Peluso, "Context effects on older consumers' cognitive age: The role of hedonic versus utilitarian goals," *Psychology & Marketing*, vol. 31, no. 2, pp. 103–114, 2014.

[31] I. Patterson and S. Pegg, "Marketing the leisure experience to baby boomers and older tourists," in *Marketing of tourism experiences*, pp. 162–180, Routledge, 2013.

[32] A. S. Grotts and T. Widner Johnson, "Millennial consumers' status consumption of handbags," *Journal of fashion marketing and management: an international journal*, vol. 17, no. 3, pp. 280–293, 2013.

[33] P. Salomoni, C. Prandi, M. Roccetti, L. Casanova, L. Marchetti, and G. Marfia, "Diegetic user interfaces for virtual environments with hmds: a user experience study with oculus rift," *Journal on Multimodal User Interfaces*, vol. 11, pp. 173–184, 2017.

[34] W. Hwang and G. Salvendy, "Number of people required for usability evaluation: the 10±2 rule," *Communications of the ACM*, vol. 53, no. 5, pp. 130–133, 2010.

[35] L. Faulkner, "Beyond the five-user assumption: Benefits of increased sample sizes in usability testing," *Behavior Research Methods, Instruments, & Computers*, vol. 35, pp. 379–383, 2003.

[36] S. Peretti, D. Tempesta, V. Socci, M. C. Pino, M. Mazza, M. Valenti, L. De Gennaro, C. Di Dio, A. Marchetti, and M. Ferrara, "The role of sleep in aesthetic perception and empathy: A mediation analysis," *Journal of Sleep Research*, vol. 28, no. 3, p. e12664, 2019.

[37] S. Mastandrea, S. Fagioli, and V. Biasi, "Art and psychological wellbeing: Linking the brain to the aesthetic emotion," *Frontiers in psychology*, vol. 10, p. 739, 2019.

[38] I. Schindler, G. Hosoya, W. Menninghaus, U. Beermann, V. Wagner, M. Eid, and K. R. Scherer, "Measuring aesthetic emotions: A review of the literature and a new assessment tool," *PloS one*, vol. 12, no. 6, p. e0178899, 2017.

[39] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

Using Virtual Reality to Assess Communicational Skills During a Collaborative Task with Time Pressure

Yasmina Kebir
Chaire Behaviour
University of Lorraine
Metz, France
email: yasmina.kebir@univ-lorraine.fr

Gaelle Nicolas
Chaire Behaviour
University of Lorraine
Metz, France
email: gaelle.nicolas@univ-lorraine.fr

Samuel Ferreira Da Silva
National Engineering School
University of Lorraine
Metz, France
email: samuel.ferreira-da-silva@univ-lorraine.fr

Pierre Chevrier
National Engineering School
University of Lorraine
Metz, France
email: pierre.chevrier@univ-lorraine.fr

Jérôme Dinet
2LPN, UR 7489
University of Lorraine
Nancy, France
email: jerome.dinet@univ-lorraine.fr

Valérie Saint-Dizier de Almeida
2LPN, UR 7489
University of Lorraine
Nancy, France
email: valerie.saint-dizier@univ-lorraine.fr

***Abstract*— Communication is an essential nontechnical skill, required in any activity involving social and professional interaction. This article presents an assessment of this specific nontechnical skill through a virtual reality (VR) simulation. The VR scenario used involve a team solving of a collective task under time pressure. We will describe the immersive environment as a tool for assessing communication skills. Twenty-three participants were included in the study and divided into four different groups. Qualitative and quantitative data were collected and will be presented to assess and compare team communication when performing a time-pressured collaborative task in an immersive environment. Three types of groups emerged, involving different and more or less effective communication.**

Keywords-virtual reality; communication; nontechnical skills; interaction; assessment; group.

I. Introduction

Communication is an essential nontechnical skill for any professional interaction. For industries, this nontechnical skill has been identified as a determining factor in the occurrence or non-occurrence of accidents in workplaces characterized by significant risks and interactions involving team working. To solve this problem, managers set out to deeply study this type of nontechnical skills, in situations analogous to those encountered in the real world of work, in order to improve safety and performance in the workspace. This study is part of a larger project undertaken by the "Behaviour" chair [1], based on a multi-faceted collaboration between researchers, industry and developers of VR scenarios. In this study, VR will be used as an innovative means to conduct communication research and analysis, providing a rich and diverse data source, a high level of realism and immersion, via a flexible and adaptable platform.

VR is used to define a computer-generated environment that can be experienced, explored, and with which a person can interact [2]. This person becomes an integral part of this virtual world via the immersion principle, and can then manipulate objects or perform a set of actions [3]. Notably, VR will be seen as an ideal means of implementing new frameworks that mimic real-life functions and situations [4] as it offers safe and flexible ways to create various environments that are easily reproducible and enables a variety of behavioral responses (e.g., language, actions, movements) to be accurately measured [5].

It has been widely demonstrated that successful teamwork depends on interaction and knowledge sharing between team members [6]. Moreover, communication has often been shown to be an important predictor of team and project performance [7]. In scientific literature, a number of studies have investigated group communication using VR [8]. However, these studies are mainly quantitative and do not highlight the qualitative aspect.

This article follows a structured outline. In Section 1, we detail the methodology employed for conducting this study. This includes an overview of our VR simulation tool, a description of the participants, and an outline of our procedural approach. Section 2 presents the overall quantitative and qualitative results obtained. Section 3 will be devoted to the concluding part, in which we will discuss our overarching findings, highlighting limitations, and outlining future prospects for our research.

II. Method

In this section, we describe the methodology used in our study. First, we describe the virtual reality scenario. We will then present the sample involved in the study. We will finish by explaining the procedure deployed.

A. Virtual reality simulation

The VR simulation involves a collective resolution of a task in an unusual environment. It reproduces an immersion in a submarine (Figure 1) in which the team must collaborate and communicate to succeed. This simulation was collaboratively developed by Virtual Rangers [9], a creation studio specializing in VR, in conjunction with researchers and industry experts.

 ISBN: 978-1-68558-154-1

Figure 1. virtual environment display.

The mission is to assemble different items produced by each participant into a final piece. The entire scenario, if successful, involves the production of 5 pieces within a given time limit. However, a number of rules must be observed including wearing the right safety gloves tailored to each item (chemical risk, explosion risk, etc.), assembling them in a particular order, and ensuring that team members choose the correct reference for each item, among other requirements.

The time pressure added by the virtual environment comes from the timer which is constantly visible to all participants. Moreover, every three errors a minute is deducted from the overall time. Another type of pressure comes from the red halo that lights up around every participant who makes a mistake. Participants are placed in an unconventional situation that generates stress through various disruptive and stressful elements.

B. Participants

A total of 23 participants aged between 21 and 25 (μ=22.5 years; σ= .97 years) were recruited to test the virtual environment. The participants were all 5th-year students from the same engineering school. They are volunteers, native speakers of French and have a strong interpersonal tie. Participants are part of the same graduating class and work on a joint project during their final year. It should also be noted that all the participants are testing this VR simulation for the first time.

C. Procedure

Participants who agreed to join the study were randomly divided into groups : three groups of six and one group of five. On arrival in the VR room, all participants consulted and signed a consent form. Next, the technical engineer described the context of the simulation, the general objective and showed them the controls to be used to perform actions in VR. During this briefing phase, they were free to ask questions to remove any ambiguity, as they had been told that during the simulation, they should and could only interact with each other. Participants were then fitted with VR headsets and hand-held controllers. A training phase is planned at the beginning of the VR so that participants can interact and familiarize themselves with the environment.

For data collection purposes, the setup incorporates a data recording function, enabling us to collect real-time verbal interaction, time spent to perform the task, video recording of their actions in the VR environment including number and type of errors.

III. RESULTS

In this third section, we detail the results obtained by analyzing our qualitative and quantitative data.

A. Quantitative DATA

The quantitative data we rely on are the time spent completing tasks and the number of errors made for each group and each part. (Figure 2).

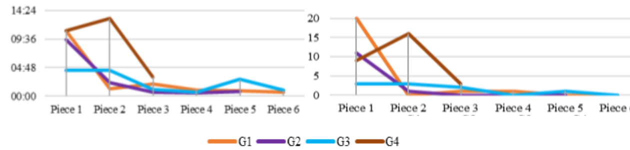

Figure 2. time taken and errors number per piece and per group.

As we can see, three types of group stand out. G1 and G2 have performances that improve considerably between piece 1 and piece 2, and stabilize from piece 3 onwards. Then, we have G3, which is relatively stable, with no significant variation in performance in the production of the 6 pieces. Finally, we have G4, whose performance fluctuates unexpectedly. Indeed, as Figure 2 shows, this group's performance is worse for piece 2 than for piece 1. The number of errors and the time taken increase for the second piece. This represents an unexpected result.

In summary, considering these quantitative data as performance indicators, G3 emerges as the top-performing group in the task, followed by G2, G1, and lastly, G4, which exhibits the highest number of errors (n=28) and a longer completion time compared to the other groups.

To better understand these variations, we will present the results of our qualitative data analysis in the next section.

B. Qualitative DATA

Analysis of qualitative data has enabled us to understand more precisely the communication that distinguishes these different groups. In general, there is a noticeable contrast in the communication dynamics employed for the production of the initial pieces compared to the final ones across all our groups. The discursive sequences produced for the construction of the initial pieces were significantly longer than those generated for the construction of the final pieces. For the initial ones, these sequences were less structured, leading to co-comprehension processes, co-construction, and sense negotiation. By contrast, the final pieces contain increasingly operative language that is concise, unambiguous, shared and subject to less interpretation. Overall, these global results are applicable to all groups. Indeed, for the production of the first piece, we observe, a poorly structured communication, with a lot of overlapping speech and long, complex sentences *"But then, I'm preparing the detonator for you, and I'm preparing the propeller, so the propeller is an S631, at risk of corrosion"*. For the last pieces, we observe the use of very short, clear and intelligible operative phrases such as: *"that's it"*, *"it's sent"*, *"ok"*. As the simulation progresses, most groups acquire increasingly precise, short and explicit communication. However, as the results of the quantitative data analysis show, G4 performed less well on piece 2 than piece 1. This unexpected result will provide us with further information on the importance of communication for this type of task.

For the G4, we observe an increase in lengthy interventions (when we compare the production of the piece 2 by the 4 groups) while for the other three groups, we observe a progressive decrease in the number of verbal interventions produced by the participants. Our analyses also revealed an important number of verbal interventions attesting a lack of involvement in achieving the team's objective: *"You think we can knock his tower down [Laughs]"*. This form of intervention, which hinders task resolution, is predominantly observed in G4, particularly during the production of the second part.

Another interactional phenomenon observed in the lowest-performing group concerns humor. In the other groups, humor was more likely to occur at the end of each phase, indicating a more relaxed atmosphere "*done, we can barbecue now"* than a lack of involvement and seriousness. Indeed, in the case of G4, humor was predominantly generated incidentally by participants, during the execution of the simulation.

Compared with other groups, help is less given when asked *"can you guys help him because I can't see what he needs to send"*. In G4, this type of intervention may go unanswered. These kinds of interactional behaviors were not observed in the other three participant groups. On several occasions, we also note that the same questions were asked several times by different participants. Compared with the other groups, G4 participants made fewer requests of the commander, and updated their situation less frequently *"wait a minute, i printed the wrong piece"*.

On the basis of our various analyses, we can conclude that this simulation enabled us to distinguish the performance levels of the groups in a fairly consistent way. The shortcomings mentioned in terms of communication skills are in line with our quantitative data. The groups with the least effective communication are characterized by a longer execution time and a higher number of errors.

IV. CONCLUSION

We are convinced that new immersive technologies, such as VR, can make a real contribution to the study of nontechnical skills, such as communication. This project, which brought together industrialists, researchers and VR experts, enabled us to study a standardized situation generating complex group dynamics. Being able to experiment and study this type of situation represents a real challenge, as part of a continuous improvement, training and learning process. Indeed, this research proposes a non-domain-specific immersive environment and investigates its role in assessing communicative skills. The vast majority of current studies investigating specific NTS do so via domain-specific VR, such as medical VR. The methodological protocol used in this study is useful both to professionals for continuing education and occupational risk reduction purposes, and to researchers for the study of different interaction situations. Future research will concentrate on a different demographic, specifically targeting professionals already employed in various industries. It would be interesting to compare the interactions/communication skills of future professionals with those who are already professionally inserted. We will also look at the debriefing process that takes place at the end of each simulation, enabling participants to engage in a process of reflexivity to improve their communication skills. For future research, the VR simulation employed can also be viewed as a method for establishing an appropriate environment conducive to smart education [10]. In this perspective, we aim measure the long-term effects of VR training on communication skills.

Further research will address some of the limitations identified in this preliminary study. Specifically, it is crucial to enhance the sample size for better generalization of the obtained results. In the course of this study, we mitigated the influence of personal relationships by deliberately choosing participants who were well-acquainted with each other and had prior collaborative experience. Moving forward, it would be prudent to consider and control for additional factors that could influence group performance, such as personality traits.

ACKNOWLEDGMENT.

We would like to thank all the students who agreed to participate in this study.

REFERENCES

[1] "La Chaire Behaviour - Chaire Behaviour." [Online]. Available: https://chaire-behaviour.com/

[2] A. K. Nassar, F. Al-Manaseer, L. M. Knowlton, and F. Tuma, "Virtual reality (VR) as a simulation modality for technical skills acquisition," *Annals of Medicine and Surgery*, vol. 71, p. 102945, Nov. 2021, doi: 10.1016/j.amsu.2021.102945.

[3] L. L. Gilson, M. T. Maynard, N. C. Jones Young, M. Vartiainen, and M. Hakonen, "Virtual Teams Research: 10 Years, 10 Themes, and 10 Opportunities," *Journal of Management*, vol. 41, no. 5, pp. 1313–1337, Jul. 2015, doi: 10.1177/0149206314559946.

[4] T. D. Parsons, A. Gaggioli, and G. Riva, "Virtual Reality for Research in Social Neuroscience," *Brain Sci*, vol. 7, no. 4, p. 42, Apr. 2017, doi: 10.3390/brainsci7040042.

[5] Z. Campbell, K. K. et al, "Utilizing Virtual Reality to Improve the Ecological Validity of Clinical Neuropsychology: An fMRI Case Study Elucidating the Neural Basis of Planning by Comparing the Tower of London with a Three-Dimensional Navigation Task," *Applied Neuropsychology*, vol. 16, no. 4, pp. 295–306, Nov. 2009, doi: 10.1080/09084280903297891.

[6] M. A. Marks, J. E. Mathieu, and S. J. Zaccaro, "A Temporally Based Framework and Taxonomy of Team Processes," *The Academy of Management Review*, vol. 26, no. 3, pp. 356–376, 2001, doi: 10.2307/259182.

[7] G. Hirst and L. Mann, "A model of R&D leadership and team communication: the relationship with project performance," *R&D Management*, vol. 34, no. 2, pp. 147–160, 2004, doi: 10.1111/j.1467-9310.2004.00330.x.

[8] J. Chen, C. Gu, J. Zhang, Z. Liu, and S. Konomi, "Sensing the Intentions to Speak in VR Group Discussions," *Sensors*, vol. 24, no. 2, Art. no. 2, 2024, doi: 10.3390/s24020362.

[9] "Virtual Rangers, creative studio for your VR & AR applications !," Virtual Rangers Luxembourg. [Online]. Available: https://virtual-rangers.com/en/

[10] N. Ye. Dmitrenko, O. V. Voloshyna, S. S. Kizim, K. V. Mnyshenko, and S. V. Nahorniak, “Smart education in the prospective teachers’ training,” *CTE Workshop Proc.*, vol. 10, pp. 414–429, Mar. 2023, doi: 10.55056/cte.568.

Describing and Predicting the Acceptability of AI and Robotics towards Professional Identity with the Revised 4-A Model

Jérôme Dinet
University of Lorraine
2LPN and Chair Behaviour
Nancy, France
Email: jerome.dinet@univ-lorraine.fr

Armand Manukyan
University of Lorraine
Nancy, France
Email: Armand.MANUKYAN@jbthiery.asso.fr

Marie Rychalski
University of Lorraine
Nancy, France
Email: rychalskimarie@gmail.com

Hirokazu Kumazaki
Nagasaki University
Nagasaki, Japan
Email: kumazaki@tiara.ocn.ne.jp

Naomi Matsuura
Elementary school attached to
Mie University Faculty of Education
Tsu, Japan
Email: matuuranaomi@yahoo.co.jp

Yuichiro Yoshikawa
Osaka University
1-3 Machikaneyama, Toyonaka,
Osaka, Japan
Email: yoshikawa@irl.sys.es.osaka-u.ac.jp

***Abstract*—There have been significant developments in social robotics in the care sector, in particular, in the fields of elder care and in the care and education of children and young people, especially those with specific disabling conditions, such as autism. Because the increasing use of digital technologies, such as artificial intelligence (AI) and robotics system may be harmful to professions and occupations, it is crucial to investigate the relationships between professional identity towards robotics systems to describe and predict the acceptability. In this communication, we present a revised version of the 4-A model (for Acceptability, Acceptation, Approval, Appropriation) to apprehend the relationships between professional identity and acceptability. The origins and the main advantages of this revised theoretical framework are presented and discussed. This communication contributes to efforts to shift the ways in which the future of work and the rise of robotics and AI are understood by proposing a new framework for articulating the resulting disruptions in relation with professional identity.**

***Index Terms*—Professional context, Robotics, Acceptability**

I. Introduction

This paper is aiming to present the revised and the more recent version of the 4-A model integrating the different components of the professional identity to better describe and predict the acceptability of technology, such as robotics systems and Artificial System. The first section is presenting the relationships between acceptability and professional identity before to discuss the links between acceptability and ethics in professional context, specially when robotics systems are used. The second section is focused on the revised A-A model, by presenting its advantages and its implication and the integration of the different components of the professional identity on acceptability.

A. Acceptability of AI and Robot in the Real Professional Context

Automation, the replacement of people in the workplace by machines is not something new, but digital technology, such as robotics systems and AI have increased the capabilities of these machines enormously. There have been significant developments in social robotics in the care sector, in particular, in the fields of elder care and in the care and education of children and young people, especially those with specific disabling conditions, such as autism. With the rapid development of technology, have humans come to regard robots as their competitors? If so, how has this perception affected human–robot interactions? [1]

The increasing use of digital technologies, such as Artificial Intelligence (AI) and robotics system may be harmful to professions and occupations. Professional role identity can be damaged as AI and robots take the place of people across a broad range of professional tasks. As increasing numbers of social robots are developed, tested and deployed, attention is shifting towards issues of user experience [or UX] – including how robots are 'accepted' by users [2] [3].

This has become both a practical and an ethical issue. On the one hand, people are probably more likely to make use of, or live with, robots if they feel comfortable with, or even like, them. On the other hand, there are important ethical issues in relation to autonomy, choice and power when it comes to introducing robots to workplaces, care settings or domestic spaces. The socially or physically vulnerable, for example, should not be coerced into interacting with robots in the place of humans. Some authors calls our attention to the potentially two-sided nature of Human-Robot Interaction (HRI). Robots can be caregivers of humans; but humans can also be the caregivers of robots [4].

The acceptability (judgement before use) of a new technology, such as a robot, could involve multiple, diverse factors. The most commonly used model to describe and predict acceptability is Nielsen's model [5], which is mainly structured around practical acceptability and usefulness. Usefulness is the degree in which a person trusts the technology to perform the desired goal, and in Nielsen's model is broken down in two further notions: Utility and Usability (Figure 1).

More recent predictive theories of technology, such as the Technology Acceptance Model (TAM; Davis [6]) or the Unified Theory of Acceptance and Use of Technology (UTAUT [7] [8] [6] [9]), are also based on a priori studies. Some more recent research in Informative Sciences, based on acceptability, focus progressively on real use and adoption [9]. The emerging theory of 'situated' acceptability proposes to consider four dimensions (individual, organizational, relational, professional/identity) of the occupational activity in the field of social psychology [10] [11], and explains how acceptability factors should be engineered by confronting a real professional context. Unfortunately, little research has been reported on acceptability – the judgment towards a product after use – where both functional and perceptive factors are studied during first use (familiarization phase). Moreover, whatever the technology considered, existing models of acceptability, essentially based on predictive methods and information sciences, were not considered relevant for the case of occupational robotics in real setting, such as educational settings or in the care sector [12] [13] [14].

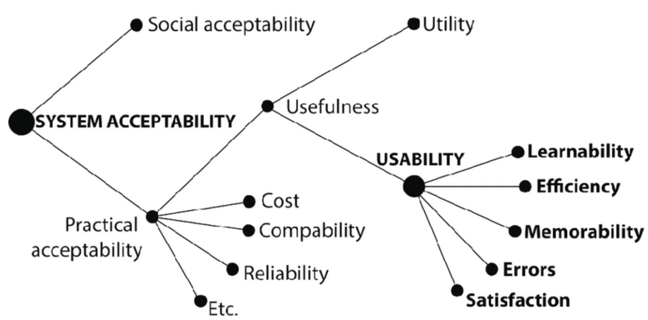

Figure 1. Nielsen's model of system acceptability.

Finally, all these existing models (e.g., Nielsen's model, TAM, UTAUT) are used to study mature and similar informative technologies, whereas innovative devices, such as robots, change the framework of acceptability through a new user-product relationship [12]. Moreover, some appropriation theories in activity ergonomics explain that it is the actual experience of the product that will influence future behavior and future adoption [15]. In addition, the acceptability of digital systems needs to consider physical and environmental aspects. This consideration of interfacing is why we believe that existing models of acceptability are not sufficiently adapted to physical user-product experience in the occupational environment.

Consequently, there is a need to involve real work situations to identify important determinants of robot acceptability, and a more holistic and usability focused approach is needed to identify obstacles to social worker acceptability that are not evident in a laboratory environment (e.g., [16]), in particular if we want to better understand the influence of robotics systems on professional identity.

B. Acceptability, Ethics and Professional Identity

How will the future world of the social care professional and education, specially for users with specific needs, evolve in this context? What will the acceptability of social robots be amongst social professionals who have different professional identities?

Professional identities refer to the way we define ourselves in relation to our work, including the values, beliefs, and practices that shape our sense of professional self. It is a complex and multifaceted concept with significant implications for both individuals and organizations. In recent years, there has been a growing interest in the study of professional identities in the healthcare and education professions (e.g., [17]). Professional identity is a crucial construct [18] that impacts many important aspects of individuals' lives such as:

- Confidence in advocating for professional opinion;
- Source of meaningfulness;
- A sense of self-worth and empowerment;
- Determination of one's moral decision-making and behavior;
- Psychological well-being.

Teachers psychological empowerment tends to be an important factor of their professional identity. It refers to teachers' confidence in their ability to do their jobs well and their belief that their work is meaningful and valuable. Teacher professional identity is seen as a sense of recognition that teachers have for the profession of teaching. In fact, Sun et al. [19] indicate that teachers with higher level of recognition of their profession will believe that their work is more meaningful and valuable. That is, they will have a higher level of psychological empowerment. Specifically, the higher the level of teachers' professional identity, the higher their psychological empowerment will be, which will lead to increased work engagement. Therefore, teachers can fully dedicate themselves to their work when they have a professional identity in terms of the profession of "teacher", which will improve their professional identity. In this way, Ding et al. [20] indicate that both psychological empowerment and professional identity were significantly and negatively related to work burnout, and psychological empowerment was significantly and positively related to professional identity.

For instance and as Figure 2 shows, a same robot and AI in a classroom can have diverse uses to improve the learning of science. In the same way, and as Figure 3 shows, to meet the needs of autistic children, the use of a robot has increased over the years. Studies [21][22][23] point to benefits in the development of academic skills and social interaction. It appears that most children with ASD prefer to interact with robots because of their simplicity, predictability and predictability. Indeed, the emotional microexpressions, behavioral variations and different voice intonations of professionals can be obstacles to understanding autistic children. However, the results of the studies cannot be systematized, as the profiles of the children and the robots used differ from one study to another. Whatever the context (Figure 2 or Figure 3), the acceptability and the use of the robotics systems are strongly to the professional identity of the teachers or the educators.

Some are objects of study for students to practice programming, others are tools which assist a teacher, some can be learning companions, and others might be autonomous teachers which provide some unit of instruction more or less in its entirely. Like most innovations, there may be a good side and a bad side, and care is needed to foster the former and counter the latter. The roles of the human teacher change over time with needs, new tools and teaching aids, but the capabilities and nature of AI promote teaching robots to new levels of relationship with the teacher and the learner as Figure 2 shows. Aids to teaching and learning are not, of course, new. Humanoid robots, however, are more active, even pro-active. Unlike the passive textbook, they can respond and adapt to each student, tailoring teaching to particular needs. There is clear evidence that they have the potential to support learning, as in teaching children about their medical conditions, developing and rehearsing learning, and testing it. Finally, robots can even do what a teacher would find difficult by his or her presence, as in teaching an ASD student while slowly accustoming that student to social interaction [24].

Identity is generally the concept that defines who a person is in relation to some phenomena, groups, objects, and social behaviors [25]. Material objects, personal characteristics, or group norms can be an integral part of identity if individuals use them to identify themselves in communities [26]. Identity has mainly been studied from two perspectives: collective and individual level. At the collective level, social identity is framed based on membership in a social group, the group's values and the culture.

Profession is one of the most important social categories [27], and professional identity is a particular form of social identity in professional settings [28]. It is 'an individual's self-definition as a member of a profession and is associated with the enactment of a professional role' [29]. As the definition suggests, enacting a particular role is an essential part of one's professional identity. This role enactment also gives rise to role identity [30] [31]—'the goals, values, beliefs, norms, interaction styles, and time horizons that are typically associated with a role' that provide a 'definition of self-in-role' [30]. Therefore, professional identity is inherently centred around professional role identity. Moreover, evolution of values, representations and interactions over time makes identity evolutionary and dynamic.

C. Professional Identity and Robotics

In this way, Appriou Ledesma [32] developed the concept of identity strategies as characteristic of a dynamic at work in adult training in France. According to Camilleri et al. [33], identity strategies are then understood as "procedures implemented (consciously or unconsciously) by an actor (individual or collective) to achieve one, or more, goals (explicitly defined or situated at the unconscious level), procedures developed as a function of the different determinations (sociohistorical, cultural, psychological) of this situation". The functioning of identity strategies thus induces a process that evolves according to the interactions experienced, the objectives pursued and the search for integration into a group, recognition (in this case professional recognition) or even self-esteem [32]. It is made up of inseparably complementary and conflicting components. It includes inherited, acquired and projected identities whose construction, in social interaction with others, generates tensions. These tensions thus lead the subject to implement identity strategies whose "objective is to safeguard the integrity of the identity, maintain the coherence of its various components, as well as guarantee the authenticity of the project of oneself for oneself (identity project)" [32]). The practical training of a professional, invested in a mission and driven by a mandate (and a professional contract), leads him or her to deploy unfixed strategies in order to exercise his or her professional identity, through precise conducts and mechanisms. Depending on one's position and relationship with others, the establishment of one's professional representation will involve strategies aimed at ensuring consistency with one's initial training or, on the contrary, at extending the shared space of common representation. Recognized as useful worldwide, Karasek's model [34] affirms the occurrence of illnesses linked to perceived stress at work and caused by potential identitary tensions [35]. He studied work-related stress in two axes: the demand (or professional constraint or workload) and the individual's control (or decision latitude or leeway) over his or her work. He hypothesizes that stress arises in work situations that high work demands (a heavy workload) and low control over them. It thus highlights the importance of assessing professionals' representation of work. To explore this idea further, Cappe et al. [36] present in their study an investigation into burnout among educators working with people with autism. The results shows the existence of increased stress in the practice of accompanying autistic people. The feeling of ineffectiveness and incompetence appears to be prevalent in the face of care difficulties. The latter is amplified among professionals who feel they have received less training than their colleagues.

Many aspects impact professional identity; Therefore, this point can be weakened by robot integration. In fact, as robots can be anthropomorphized, a cognitive bias can appear, such as social comparison. Anthropomorphism is assigning human-like traits, emotions, and behaviors to non-human entities. As a result, the perception of self-worth, confidence, and psychological well-being are impacted because employee's comparison implies that robots can replace themselves. Robot anthropomorphism can influence employees' perception of their job insecurity in work situations. This feeling of insecurity is sometimes created by employees' comparison due to anthropomorphic thinking and can impact professional identity in work situations.

Not surprisingly, different viewpoints exist across culture [37]. Moreover very few authors have investigated the relationships between professional identities and acceptability of robotics systems [38]. For instance:

- Cahill et al. [39] highlighted that available technology had been successfully integrated into the care plans of patients in Ireland, but caregivers perceived it to be prohibitively expensive;
- Wolbring and Yumakulov [40] reveal that staff in a Canadian disability organisation are content to work with social robots as long as they perform repetitive tasks that:

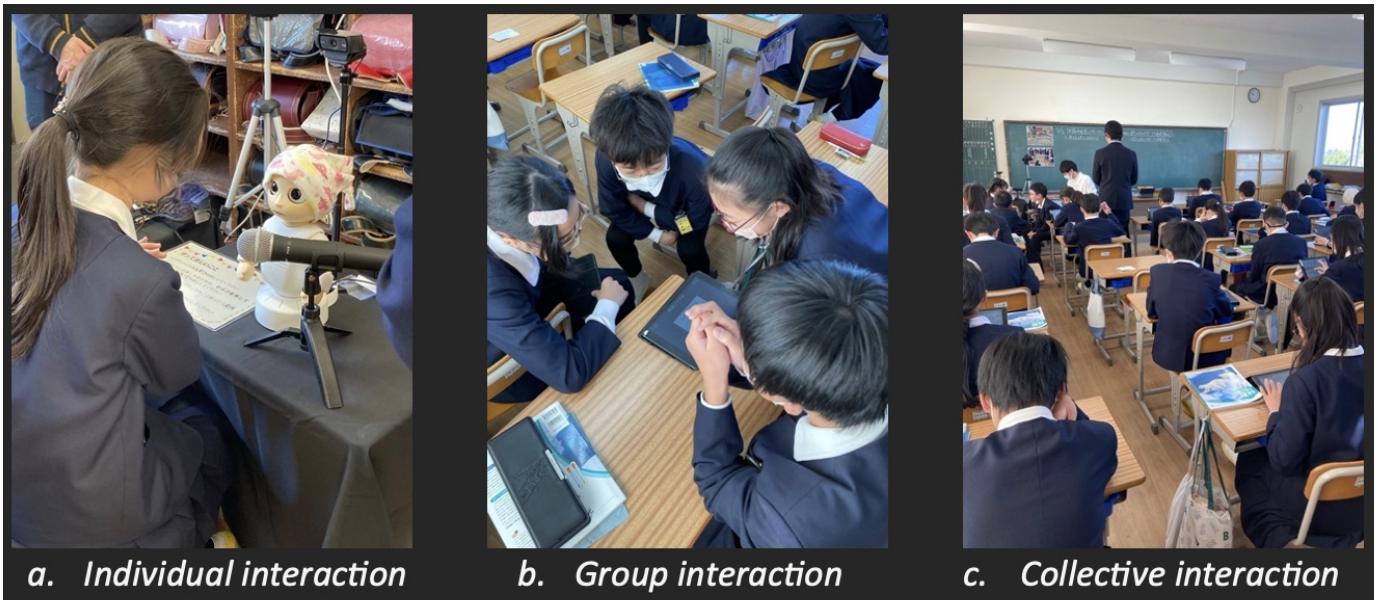

Figure 2. The use of a robot and IA to improve science learning in a classroom: the different contexts from personal, group and collective interactions [25] [26]

Figure 3. The use of a robot with a young child with Autism Spectrum Disorder (ASD)

"did not require mimicking human interaction and touch" (p. 465);
- Conti et al. [41] provide insights into the acceptability of robots in the education and care of children in Italy, uncovering that established practitioners are largely skeptical of such innovations, while less experienced degree students in psychology and education demonstrate a "significantly higher willingness to use" robots. Pragmatically, they find that "intention to use" a (hypothetical) robot is "mainly predicted by the perception that it will enhance and facilitate the educational process". Moreover, they report that "practitioners have a clearer view than students of the educational and therapeutic tools available and their effectiveness. They can easily identify the current technology difficulties and limitations" (including cost).

Working with, alongside or even for robots will have significant implications for social professional practice and identity. Practitioners may benefit from the opportunity to engage with and, if deemed appropriate, develop the skills required to work in collaboration with social robots. Those involved in the education and formation of the social professionals of the future have an obligation to stimulate and facilitate debate that may, as a parallel outcome, lead to debates about the broader philosophical, ethical, social and practical nature of 'care' itself.

D. Technology and Professional/Occupational Role Identity Change

Among the various drivers of social change, technology has long been considered an essential factor in professional settings [42]. It has recently become still more vital due to the increasing impact of digital technology on professions and occupations [43] [44]. However, as Goto says [27], studies on professional and occupational role identity have rarely investigated the impact of technology.

New technology does not enter an occupational field fully defined but is constituted within the context [45] [46] [45]. As such, technology has a way of influencing professional and occupational identity through a peculiar mechanism. Past studies have highlighted three important aspects of this mechanism.

- individual-/group-level studies have revealed that new technology itself can trigger professional and occupational identity reconfiguration and give rise to a new identity through professionals' new practices and boundary negotiations with others;
- Very few researchers have addressed the collective-level identity shift;
- Only some studies have implied that the implementation of new technology, such as robot among professionals may have an important link with the shift of professional identity;

II. The Revised 4-A Model: Acceptability, Acceptation, Approval, Appropriation

As Figure 4 shows, the actual 4-A model based on [26] [25] is an innovative model providing an explanation of the temporal process of appropriation of a digital device, such as a robot (for a complete presentation of the model, see [25] [26]. Emerging technology is not an identity threat per-se, and the relation between human and robot, regardless the professional identity, need to clarify the dependence between these two partners (either partnership, or master-slave).

A. Origins of the 4-A Model

Several studies related to the TAM theory [47] [48] [49] or the UTAUT theory [7] [8] [6] describe the role of professional identity on future acceptability and acceptance of digital devices [47] [48] [49]. But even if all these prior studies related to TAM or UTAUT theories provide very interesting results, they have four important limitations that prevent to generalize results:

- Data are often collected by using questionnaires and surveys, i.e., only attitudes, opinions and verbalization are collected;
- Data are often collected during only one-shot setting, and thus do not investigate the longitudinal and temporal process of appropriation across the time;
- They assume that the effective use of a digital device means that this device is accepted;
- Professional context and environments (physical and social) are rarely considered.

It is the reason why a new model has been created (called 4-A for "Acceptability, Acceptation, Adoption, and Appropriation") to better describe and predict the complex processes involved from the acceptability to the effective use of digital technology and to better understand the relationships with the professional identity.

B. The 4Model: its Advantages and its Implications

The 4-A model has several advantages:

- This 4-A model allows to better understand the relationships between attitudes, opinions and effective behaviours;
- If attitudes can determine behaviours (as other theoretical frameworks argue), the 44 model states that behaviours can influence attitudes by retro-feedback;
- In the 4-A model, the temporal and longitudinal dimensions related to the appropriation are included by distinguishing before and after the implementation of the device in the context. So, dynamics of the human behaviours is crucial in the 4-A model, by considering that attitudes and behaviours can change across the time;
- In the same way, there is a remarkable amount of variation in the beliefs, attitudes, professional identity and values held by people around the world. These views are often cultural, meaning that they are, at least to some extent, socially learned and socially transmitted. They

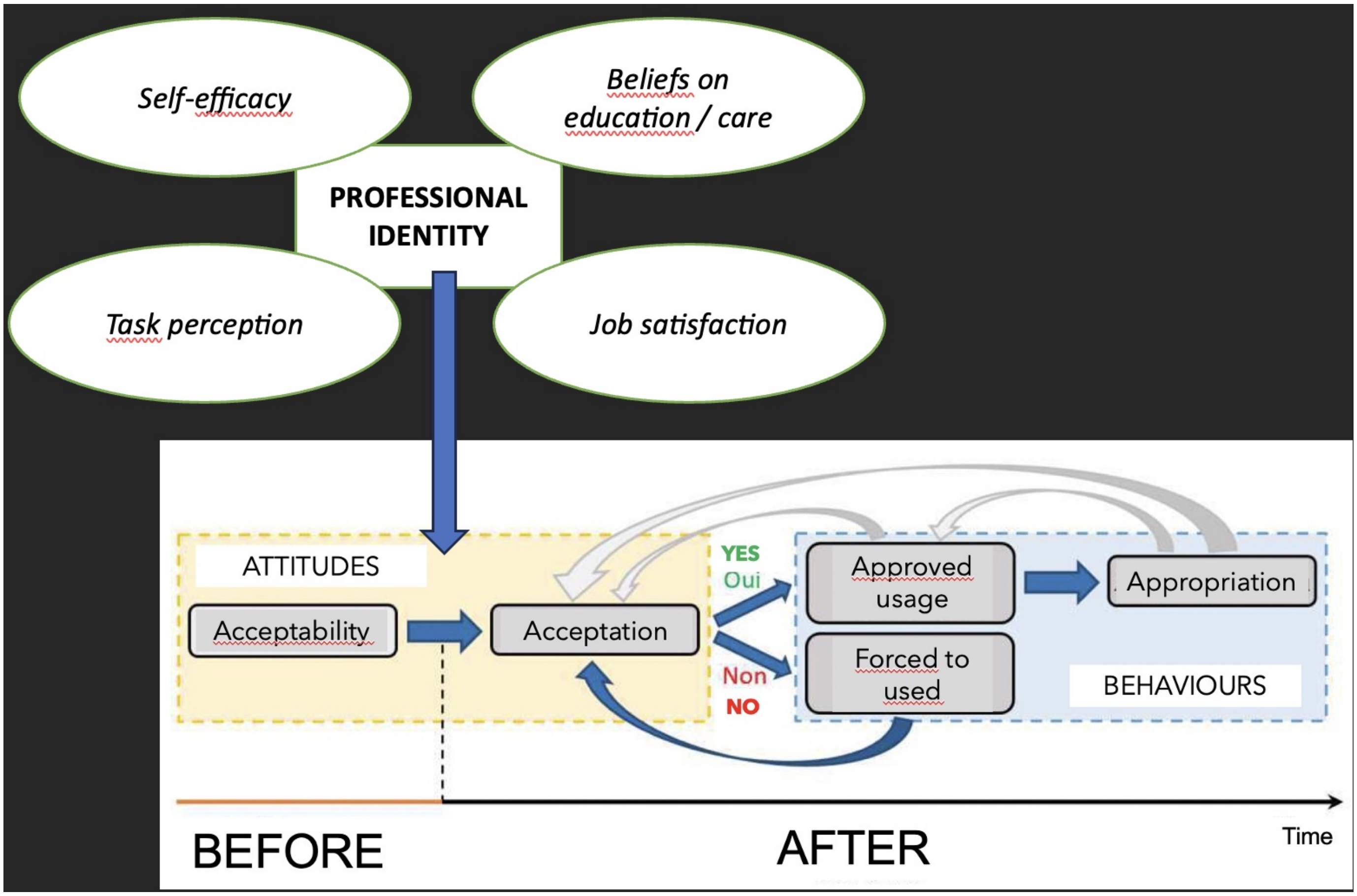

Figure 4. The revised version of the 4-A Model [25] [26]

are often shaped by tradition; namely, this transmission and persistence of cultural values across generations are captured by the 4-A model;

- The use of a device, such as a robot, does not necessarily mean that this device is approved and accepted because individual can be forced to use the device. It is the reason why two types of use are distinguished in the 4-A model: Approved use (i.e., where individual is agree to use freely and.or s/he can be convinced) versus Forced use (i.e., where individual is obliged to use the device for instance, by his/her hierarchy). In other words, according to the 4-A model, an effective use of a device does not necessarily mean that this device is accepted: in some cases, the use is forced and thus, does not indicate that the device is really accepted;

This 4-A model is the only one model that considers representations, cognitive biases, as well as the tool's ease of use and adaptability, offering insights into the integration process. This model is also interesting from an ecological point of view by its consideration of professional's perceptions of robots and their interaction with them. The 4-A model highlights that the acceptance of the tool impacts its adoption and incorporation. Hence, professional's view of the robot, its ease of use and the associated usage-related challenges serve as perspective factors for its practical utilization. A progressive handling of the tool allows to facilitate teachers' comprehension and to focus on the use to offer an efficient support, with less workload for professionals.

C. The Integration of Professional Identity into the Revised 4-A model

The revised and the more recent version of the 4-A model integrates the different components of the professional identity to better describe and predict the acceptability of technology, such as robotics systems (Figure 3).

There is consensus in the research that professional identity is a multidimensional concept, but still no unanimous agreement on its central components [50] [51] [52]. However, four main components can be identified that have emerged from a variety of studies as manifestations of professional identity in teacher educators:

- The first of these is task perception, i.e., the individual understanding of the tasks for which a person feels responsible;
- The second is self-efficacy, the perception of one's ability to deal successfully with the specific requirements of one's profession;

 ISBN: 978-1-68558-154-1

- The third component is the perception of satisfaction (or failure), since experiencing success in a job may lead to a feeling of satisfaction, whereas the experience of failure may result in a feeling of stress;
- The fourth component of professional identity is the personal system of beliefs on teaching and how to put them into practice (in healthcare or in education for instance).

The four aforementioned components of professional identity are important for actions and behaviour in the workplace and may therefore influence the individual's performance, the quality of their actions and their attitudes (e.g., [53]. In other words, these four components of professional identity influence directly attitudes towards technology and then they influence acceptability and thus, the next steps of the process (acceptation, approval, appropriation).

III. Discussion

Robots have become increasingly embedded in the very core of many firms' products, services, and operations, which implies that people's roles and relationships become somewhat inseparable from their interactions with technology and in changing professional roles, which influences one's occupational identity [54] [55] [56] [57]. In particular in the fields of elder care and in the care and education of children and young people, especially those with specific disabling conditions, such as autism, this increasing use of robotics systems and Artificial Intelligence (AI) may be harmful to professions and occupations and some authors have investigated the disruptive potential of robotics [58].

The real and imagined disruptions of increasingly automated work that will unfold over the coming decades will have profound implications. From the everyday experiences of individual value and worth to the priorities of federal legislation and resource allocation, the reconfiguration of work will have widespread impact. This communication contributes to efforts to shift the ways in which the future of work and the rise of robotics and AI are understood. By proposing a new framework for articulating the resulting disruptions in relation with professional identity, our communication aims to engage with a range of discussions around researches, policy priorities, legal frameworks, and stakeholder decision-making processes. In other words, the crucial questions are: What of the humans who currently provide human-to-human social and educational care? What of their future professional training and identity needs in a world of care and education provision delivered by or, at the very least, augmented by AI and robots?

IV. Conclusion and Future Works

This paper aimed to present the revised and the more recent version of the 4-A model integrating the different components of the professional identity to better describe and predict the acceptability of technology, such as robotics systems. Actually, this revised 4-A model is the only one model that considers representations, cognitive biases, as well as the tool's ease of use and adaptability, offering insights into the integration process. Because the revised 4-A model highlights that the acceptance of the tool impacts its adoption and incorporation, it is also interesting from an ecological point of view by its consideration of professional's perceptions of robots and their interaction with them.

Emerging technology is not an identity threat per-se. Mainly, it depends upon how the professional appraises and evaluates it against the current definition of identity [59] [60]. But technology can be considered as disruptive if it fundamentally displaces an earlier technology, forces organizations to fundamentally change their business model or leads to radical organizational change [61]. Currently, Artificial Intelligence (AI) and robotics systems are considered as two major disruptive technologies in healthcare and education.

Even if the use of robots in workplaces in healthcare or education care can offer multiple advantages, professional role identity can be damaged as AI and robots take the place of people across a broad range of professional tasks. For that reason, professional identity can be managed with specific goals of using robots in work situation and limits of their use had to be explain [62]. As the implementation of robot aims to alleviate mental and physical limits [63], specific tasks must be given to robots, like repetitive or tiresome works, to facilitate the acceptability and to preserve professional identity. In fact, out of place disruption creates negative effects on social perception of the user during a task or, on the willingness to work in collaboration and impacts the HRI [64] [65].

Note that this research was partially supported by JST-Mirai Program Grant Number JPMJMI22J3 and Mie Prefecture school (Japan) and the Association Jean-Baptiste Thiery in Maxeville (France)

References

[1] D. Yang and X. He, "The transition of robot identity from partner to competitor and its implications for human–robot interaction," *International Journal of Social Robotics*, vol. 14, no. 9, pp. 2029–2044, 2022.

[2] C. Bartneck and M. Keijsers, "The morality of abusing a robot," *Paladyn, Journal of Behavioral Robotics*, vol. 11, no. 1, pp. 271–283, 2020.

[3] S. Olatunji, T. Oron-Gilad, V. Sarne-Fleischmann, and Y. Edan, "User-centered feedback design in person-following robots for older adults," *Paladyn, Journal of Behavioral Robotics*, vol. 11, no. 1, pp. 86–103, 2020.

[4] M. Boden, J. Bryson, D. Caldwell, K. Dautenhahn, L. Edwards, S. Kember, P. Newman, V. Parry, G. Pegman, T. Rodden, *et al.*, "Principles of robotics: regulating robots in the real world," *Connection Science*, vol. 29, no. 2, pp. 124–129, 2017.

[5] J. Nielsen, "Estimating the number of subjects needed for a thinking aloud test," *International journal of human-computer studies*, vol. 41, no. 3, pp. 385–397, 1994.

[6] V. Venkatesh, F. Davis, and M. G. Morris, "Dead or alive? the development, trajectory and future of technology adoption research," *The Development, Trajectory and Future of Technology Adoption Research (April 27, 2007). Venkatesh, V., Davis, FD, and Morris, MG "Dead or Alive*, pp. 267–286, 2007.

[7] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Management science*, vol. 46, no. 2, pp. 186–204, 2000.

[8] V. Venkatesh, "Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model," *Information systems research*, vol. 11, no. 4, pp. 342–365, 2000.

[9] V. Venkatesh, J. Y. Thong, and X. Xu, “Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology,” *MIS quarterly*, pp. 157–178, 2012.
[10] M.-E. B. Chaumon, “L’acceptation située des technologies dans et par l’activité: premiers étayages pour une clinique de l’usage,” *Psychologie du Travail et des Organisations*, vol. 22, no. 1, pp. 4–21, 2016.
[11] H. Anderson, *Professional identity and the advanced nurse practitioner in primary care: a qualitative study*. PhD thesis, University of York, 2017.
[12] M. Akrich, “La construction d’un système socio-technique. esquisse pour une anthropologie des techniques,” *Anthropologie et sociétés*, vol. 13, no. 2, pp. 31–54, 1989.
[13] A. Moyon, E. Poirson, and J.-F. Petiot, “Development of an acceptance model for occupational exoskeletons and application for a passive upper limb device,” *IISE Transactions on Occupational Ergonomics and Human Factors*, vol. 7, no. 3-4, pp. 291–301, 2019.
[14] R. Hensel and M. Keil, “Subjective evaluation of a passive industrial exoskeleton for lower-back support: A field study in the automotive sector,” *IISE Transactions on Occupational Ergonomics and Human Factors*, vol. 7, no. 3-4, pp. 213–221, 2019.
[15] P. Rabardel, *Les hommes et les technologies; approche cognitive des instruments contemporains*. Armand colin, 1995.
[16] S. Spada, L. Ghibaudo, S. Gilotta, L. Gastaldi, and M. P. Cavatorta, “Investigation into the applicability of a passive upper-limb exoskeleton in automotive industry,” *Procedia Manufacturing*, vol. 11, pp. 1255–1262, 2017.
[17] M. Cornett, C. Palermo, and S. Ash, “Professional identity research in the health professions—a scoping review,” *Advances in Health Sciences Education*, vol. 28, no. 2, pp. 589–642, 2023.
[18] A. Fitzgerald, “Professional identity: A concept analysis,” in *Nursing forum*, vol. 55, pp. 447–472, Wiley Online Library, 2020.
[19] B. Sun, F. Zhu, S. Lin, J. Sun, Y. Wu, and W. Xiao, “How is professional identity associated with teacher career satisfaction? a cross-sectional design to test the multiple mediating roles of psychological empowerment and work engagement,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 15, p. 9009, 2022.
[20] J. Ding and Z. Xie, “Psychological empowerment and work burnout among rural teachers: Professional identity as a mediator,” *Social Behavior and Personality: an international journal*, vol. 49, no. 6, pp. 1–9, 2021.
[21] C. Huijnen, M. Lexis, and L. De Witte, “Robots as new tools in therapy and education for children with autism,” *International Journal of Neurorehabilitation*, vol. 4, no. 4, pp. 1–4, 2017.
[22] F. Sartorato, L. Przybylowski, and D. K. Sarko, “Improving therapeutic outcomes in autism spectrum disorders : Enhancing social communication and sensory processing through the use of interactive robots.,” *Journal of Psychiatric Research*, vol. 90, no. 2, pp. 1–11, 2017.
[23] P. Pennisi, A. Tonacci, G. Tartarisco, L. Billeci, L. Ruta, S. Gangemi, and G. Pioggia, “Autism and social robotics : A systematic review.,” *Autism Research*, vol. 9, no. 2, pp. 165–183, 2016.
[24] D. P. Newton and L. D. Newton, “Humanoid robots as teachers and a proposed code of practice,” in *Frontiers in education*, vol. 4, p. 125, Frontiers Media SA, 2019.
[25] C. Bauchet, B. Hubert, and J. Dinet, “From acceptability of digital change to appropriation of technology: The 4a model,” in *17th EARA Conference” Adolescence in a rapidly changing world”*, 2020.
[26] C. Bauchet, B. Hubert, and J. Dinet, “Entre acceptabilité et appropriation des outils numériques intégrés dans le système éducatif: Le modèle des 4a,” in *13ème colloque international RIPSYDEVE La psychologie du développement et de l’éducation pour le 21ème siècle: nouveaux objets, espaces et temporalités*, pp. 158–161, 2020.
[27] M. Goto, “Collective professional role identity in the age of artificial intelligence,” *Journal of Professions and Organization*, vol. 8, no. 1, pp. 86–107, 2021.
[28] Y. Kyratsis, R. Atun, N. Phillips, P. Tracey, and G. George, “Health systems in transition: Professional identity work in the context of shifting institutional logics,” *Academy of Management Journal*, vol. 60, no. 2, pp. 610–641, 2017.
[29] S. Chreim, B. E. Williams, and C. Hinings, “Interlevel influences on the reconstruction of professional role identity,” *Academy of management Journal*, vol. 50, no. 6, pp. 1515–1539, 2007.
[30] B. Ashforth, *Role transitions in organizational life: An identity-based perspective*. Routledge, 2000.
[31] E. Goodrick and T. Reay, “Florence nightingale endures: Legitimizing a new professional role identity,” *Journal of Management studies*, vol. 47, no. 1, pp. 55–84, 2010.
[32] L. Appriou-Ledesma, *Le sentiment identitaire professionnel*. PhD thesis, CNAM, 2018.
[33] C. Camilleri, L. E.-M. Kastersztein, Joseph, and T.-L. I. . V. A. Malewska-Peyre, Hanna, *Stratégies identitaires*. Presses Universitaires de France, 1992.
[34] R. Karasek, “Occupational distribution of psychological demands and decision latitudes,” *International Journal of Health Services*, vol. 19, no. 3, pp. 481–508, 1989.
[35] F. Chapelle, *Modèle de Karasek.*, ch. 16, pp. 107–112. Dunod, 2018.
[36] Cappe, M.-C. Rougé, and E. Boujut, “Burnout des professionnels de l’éducation spécialisée intervenant auprès d’individus ayant un trouble du spectre de l’autisme : Rôle des antécédents psychosociaux et des processus transactionnels.,” *Psychologie du Travail et des Organisations*, vol. 2, pp. 125–148, 2015.
[37] M. M. De Graaf, S. Ben Allouch, and J. A. van Dijk, “Long-term evaluation of a social robot in real homes,” *Interaction studies*, vol. 17, no. 3, pp. 462–491, 2016.
[38] P. Share and J. Pender, “Preparing for a robot future? social professions, social robotics and the challenges ahead,” *Irish Journal of Applied Social Studies*, vol. 18, no. 1, p. 4, 2018.
[39] S. Cahill, E. Begley, J. P. Faulkner, and I. Hagen, ““it gives me a sense of independence”–findings from ireland on the use and usefulness of assistive technology for people with dementia,” *Technology and Disability*, vol. 19, no. 2-3, pp. 133–142, 2007.
[40] G. Wolbring and S. Yumakulov, “Social robots: views of staff of a disability service organization,” *International journal of social robotics*, vol. 6, pp. 457–468, 2014.
[41] D. Conti, S. Di Nuovo, S. Buono, and A. Di Nuovo, “Robots in education and care of children with developmental disabilities: a study on acceptance by experienced and future professionals,” *International Journal of Social Robotics*, vol. 9, pp. 51–62, 2017.
[42] M. Noordegraaf, “Risky business: How professionals and professional fields (must) deal with organizational issues,” *Organization studies*, vol. 32, no. 10, pp. 1349–1371, 2011.
[43] M. Smets, T. Morris, A. Von Nordenflycht, and D. M. Brock, “25 years since ‘p2’: Taking stock and charting the future of professional firms,” *Journal of Professions and Organization*, vol. 4, no. 2, pp. 91–111, 2017.
[44] B. Hinings, T. Gegenhuber, and R. Greenwood, “Digital innovation and transformation: An institutional perspective,” *Information and Organization*, vol. 28, no. 1, pp. 52–61, 2018.
[45] J. R. Zetka Jr, “Occupational divisions of labor and their technology politics: The case of surgical scopes and gastrointestinal medicine,” *Social Forces*, vol. 79, no. 4, pp. 1495–1520, 2001.
[46] P. M. Leonardi and S. R. Barley, “What’s under construction here? social action, materiality, and power in constructivist studies of technology and organizing,” *The Academy of Management Annals*, vol. 4, no. 1, pp. 1–51, 2010.
[47] F. D. Davis, “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” *MIS quarterly*, pp. 319–340, 1989.
[48] F. D. Davis, “User acceptance of information technology: system characteristics, user perceptions and behavioral impacts,” *International journal of man-machine studies*, vol. 38, no. 3, pp. 475–487, 1993.
[49] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, “User acceptance of computer technology: A comparison of two theoretical models,” *Management science*, vol. 35, no. 8, pp. 982–1003, 1989.
[50] E. T. Canrinus, M. Helms-Lorenz, D. Beijaard, J. Buitink, and A. Hofman, “Self-efficacy, job satisfaction, motivation and commitment: Exploring the relationships between indicators of teachers’ professional identity,” *European journal of psychology of education*, vol. 27, pp. 115–132, 2012.
[51] G. Kelchtermans, “Who i am in how i teach is the message: self-understanding, vulnerability and reflection,” *Teachers and Teaching: theory and practice*, vol. 15, no. 2, pp. 257–272, 2009.
[52] E. Richter, M. Brunner, and D. Richter, “Teacher educators’ task perception and its relationship to professional identity and teaching practice,” *Teaching and Teacher Education*, vol. 101, p. 103303, 2021.
[53] M. Kunter, U. Klusmann, J. Baumert, D. Richter, T. Voss, and A. Hachfeld, “Professional competence of teachers: Effects on instructional quality and student development.,” *Journal of educational psychology*, vol. 105, no. 3, p. 805, 2013.

[54] M. Carter, S. Petter, V. Grover, and J. B. Thatcher, “It identity: a measure and empirical investigation of its utility to is research,” *Journal of the Association for Information Systems*, vol. 21, no. 5, p. 2, 2020.

[55] F. Pemer, “Enacting professional service work in times of digitalization and potential disruption,” *Journal of Service Research*, vol. 24, no. 2, pp. 249–268, 2021.

[56] J. Wirtz, P. G. Patterson, W. H. Kunz, T. Gruber, V. N. Lu, S. Paluch, and A. Martins, “Brave new world: service robots in the frontline,” *Journal of Service Management*, vol. 29, no. 5, pp. 907–931, 2018.

[57] E. Smailhodzic and A. M. Nijgh, “(how) does occupational identity change due to co-working with robots?,” 2022.

[58] P. Moradi and K. Levy, “The future of work in the age of ai: Displacement or risk-shifting?,” 2020.

[59] M.-K. Stein, R. D. Galliers, and M. L. Markus, “Towards an understanding of identity and technology in the workplace,” *Journal of Information Technology*, vol. 28, no. 3, pp. 167–182, 2013.

[60] M.-K. Stein, S. Newell, E. L. Wagner, and R. D. Galliers, “Coping with information technology,” *Mis Quarterly*, vol. 39, no. 2, pp. 367–392, 2015.

[61] E. Jussupow, K. Spohrer, A. Heinzl, and C. Link, “I am; we are-conceptualizing professional identity threats from information technology,” 2018.

[62] M. L. Schrum, G. Neville, M. Johnson, N. Moorman, R. Paleja, K. M. Feigh, and M. C. Gombolay, “Effects of social factors and team dynamics on adoption of collaborative robot autonomy,” in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 149–157, 2021.

[63] F. A. Storm, M. Chiappini, C. Dei, C. Piazza, E. André, N. Reißner, I. Brdar, A. Delle Fave, P. Gebhard, M. Malosio, *et al.*, “Physical and mental well-being of cobot workers: A scoping review using the software-hardware-environment-liveware-liveware-organization model,” *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 32, no. 5, pp. 419–435, 2022.

[64] M. Natarajan, E. Seraj, B. Altundas, R. Paleja, S. Ye, L. Chen, R. Jensen, K. C. Chang, and M. Gombolay, “Human-robot teaming: grand challenges,” *Current Robotics Reports*, vol. 4, no. 3, pp. 81–100, 2023.

[65] E. Roesler, M. Vollmann, D. Manzey, and L. Onnasch, “The dynamics of human–robot trust attitude and behavior—exploring the effects of anthropomorphism and type of failure,” *Computers in Human Behavior*, vol. 150, p. 108008, 2024.

Work-In-Progress - The Impact of Virtual Reality on Pain Management During Orthodontic Debonding

Ines Bouhlal
Reims hospital, URCA
Reims, France
Email: ines.bouhlal@gmail.com

Jerome Dinet
Université de Lorraine
Nancy, France
Email: jerome.dinet@univ-lorraine.fr

Aurélie Mailloux
Reims hospital, URCA
Reims, France
Email: aurelie.mailloux@univ-reims.fr

***Abstract*—Several previous studies reported that the use of Virtual Reality (VR) can reduce patient anxiety and pain, and can decrease the need of analgesics during medical care. Although patients may experience pain when their fixed orthodontic appliance is removed, this topic is poorly documented, as evidenced by the scarcity of publications. This paper is aiming to present a Work-In-Progress about an experiment investigating the impact of a VR system on pain and anxiety for patients during their orthodontic appliance debonding, with an experiment conducted with more than 100 adolescents and adults. Moreover, we investigate the correlation between patients' anxiety before debonding and their perception of pain, by comparing with other techniques to manage pain. To our knowledge, VR has never been tested during orthodontic procedures and our experiment is the first one investigating the benefits of VR for patients during orthodontic debonding.**

***Keywords*—Orthodontic; Debonding; Virtual Reality; Pain; Anxiety**

I. Introduction

Some orthodontic procedures, such as the application of an orthodontic force (e.g., archwire, elastomeric chain), the installation of mini screws and the appliance debonding involve discomfort and/or pain [1]. Pain is recognized to have a negative impact on physical, psychological and social dimensions of quality of life. Although numerous treatments are available for pain reduction, people suffering moderate to severe pain are often unable to find adequate pain relief and this has led to a great interest in finding novel strategies to reduce pain [2].

Because Virtual Reality (VR) distraction techniques, as a non-invasive technique, seems to be promising non-pharmacological approach for pain management [3][4][5], VR has become popular in clinical research studies as an innovative distractor technique and is a promising technology to enhance dental education [6][7] and to help patients [8][9].

However, there are very few studies dedicated to the effectiveness of VR on pain management for dentistry patients and more specifically during orthodontic debonding [10][11]. So, the purpose of this paper is to present a Work-In-Progress about an experiment investigating the impact of a VR system on pain and anxiety for patients during their orthodontic appliance debonding.

The remainder of this article is organized as follows. Section II describes the context on VR and pain management. This is followed by the method applied to assess the efficiency of VR on pain management during the removal of fixed orthodontic appliance (in section III). Section IV presents the progresses of the study, and outlines several elements of future research to be conducted on the subject.

II. Context

This section gives an overview over the patients' pain felt during their orthodontic treatment, especially during the appliance debonding procedure (Subsection II-A). Subsection II-B describes the existing methods to manage patients' pain during orthodontic treatment (e.g., communication, video, music, finger pressure). The interests of using VR as an innovative means to decrease patients' pain and anxiety during orthodontic debonding are described in Subsection II-C. Finally, Subsections II-D and II-E outline the goal of this study and its ethical aspects, respectively.

A. Pain during Orthodontic Treatment

The orthodontic bracket bonding interface, between bracket and tooth, should be powerful to endure forces during two years (i.e., the average orthodontic treatment duration) and weak enough to allowed a comfortable, harmless, and quick removal [12][13]. However, the second requirement does not seem to be accomplished [13].

As Figure 1 shows, the current method to remove a metal orthodontic bracket is to position the blades of the ligature cutting pliers between the base and the dental surface [13]. The pain and discomfort felt during this procedure depend on, amongst others, the force required to remove the bracket.

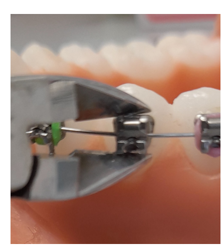

Fig. 1. Metal bracket debonding using a cutting ligature plier

Orthodontic current procedures involve pain and discomfort. Indeed, 95 percents of orthodontic patients experienced various levels of pain or discomfort during their orthodontic

 ISBN: 978-1-68558-154-1

treatment, including the removal of their appliance [12][14–20], orthodontic pain resulting from ischemia, inflammation, pressure and oedema of the periodontium [17]. The required forces to remove fixed appliance increase the pressure and ischemia of the periodontal ligament.

Several factors have an influence on pain for patients during orthodontic procedures, such as the area concerned by the intervention [21][22][23] or the bracket material (ceramic or metal) [21].

Moreover, because pain is a subjective response with important individual variations, various factors, such as gender [1][24], individual pain threshold [18][20], motivation, cultural differences, and previous negative dental experience [16], could influence the level of pain perception (and modulate pain expression) of the patient [24].

B. Current Methods to Manage Pain During Debonding

Several methods have been already tested to minimize pain during debonding, like using different orthodontic pliers, laser application, analgesics, ultrasound, biting cotton roll, practitioner finger pressure or occlusal bite wafers [1][16][21][24][25].

But there is weak evidence indicating using the (i) debonding dedicated tool rather than the most common instrument (ligature cutting plier), (ii) laser, (iii) ultrasounds during debonding reduce patients' pain perception [22]. Even if some studies tend to show that premedication can have a positive impact on pain during debonding [1], alternative pain management approaches are essential to consider [26].

Neutral statements, positive suggestions, or providing distraction are essential to improve patient treatment experience [27][28]. The practitioner's chosen words during care have a significant influence on the experience of pain. Moreover, clear communication decrease patients' anxiety [29]. Conversely, an inadequate communication can affect negatively the pain (i.e., as a *nocebo* effect) [27].

C. VR as an Innovative Technique to Manage Pain

Recent theories of pain highlighted that sensory stimuli from the environment can also influence the patient pain experience [30][31]. For example, in conscious patients undergoing surgery, watching a movie and listening to music can significantly reduce anxiety with equivalent results [32]. In the same way, VR can have a positive impact on pain by modulating the activation of some brain zones, including the anterior cingulate cortex, insula and amygdala, which are involved in the emotional and attentional pain pathways [33]. VR is efficient in reducing anxiety, decreasing pain levels, and improving patient satisfaction, and several studies have demonstrated that VR have positive effects for the care of several diseases and patient of different ages, such as burned patients [34], patients with breast cancer [35], children with cancer [36][37], children with tooth decay [38][39], elderly patients with cancer [40].

There are some years ago, VR was investigated exclusively for relaxation sessions before care and especially general anesthesia [41], all the systems being expensive and oversized. Today, VR headsets are henceforth smaller, comfortable and easy to use, usable during dental care as well [29].

D. Objectives of the Study

The current goal of this paper is to present a Work-In-Progress performed to investigate positive benefits on patient pain perception by using VR distraction during orthodontic debonding. With an experiment conducted on more than 100 adolescents and adults, this study investigates the benefits of VR on patient pain perception depending on the appliances type (i.e., ceramic or metal brackets), and the correlation between patients' anxiety before debonding and their perception of pain, by comparing with other techniques to manage pain.

E. Ethics

This study has been registered in the French General Register of Data Protection before commencement of the study. A non-opposition document was given to patients and/or the parents prior to participation.

III. Method

In this section, the chosen methodology to assess the impact of VR on patients' pain and anxiety during the debonding of their orthodontic appliance is described and justified (i.e., sample, VR headset, debonding procedure and scales).

A. Participants

All participants are patients (over 12 years-old) undergoing orthodontic treatment with a debonding of their fixed appliance scheduled between September 2023 and December 2024. They are distributed in two groups: VR group vs No-VR group. Included patients are all taken care of in our department or in the liberal practice, had benefited from a mandibular and maxillary orthodontic treatment by bonded metal or ceramic brackets (GC self-ligating brackets or AO non self-ligating brackets). Several criterias have been used (non inclusion: Cognitive disorder; Non-French understanding; Medical history of epilepsy; Blindness; Refusal, Medical history of taking analgesic or anxiolytic in the last 24h; Incomplete appliance at the time of debonding; Active periodontal disease with tooth mobility; Craniofacial dysmorphology; Presence of miniscrews).

B. Design and the Independent Factor

From early May 2024, VR distraction during debonding will be offered to patients as a new standard care protocol. To evaluate this practice, we carry out an observational before–after multicentric study (i.e., Hospital of Reims and Association Dentaire pour l'Enfance of Maisons-Alfort). We compare demographic data (gender and age), orthodontic treatment characteristics, anxiety and pain scores (general and by sextant) of 122 patients during the debonding of their orthodontic fixed appliance with or without the us of VR. The removal of the orthodontic appliance must take place from September 2023 to december 2024, over the two periods defined as "before" (called the No-VR group) and "after" (called the VR group) implementing the VR protocol.

 ISBN: 978-1-68558-154-1

To calculate the number of subjects required, we used the Wilcoxon–Mann–Whitney test, with a power of 90 percents, an alpha risk at 0.05 percent and 15 percents to make up for missing data. As Figure 2 shows, the inclusion of 122 patients (33 per group) was necessary to show a statistically significant difference on the primary objective.

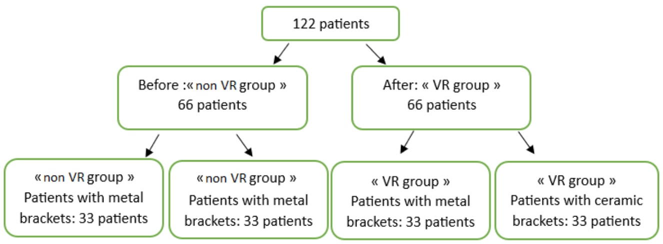

Fig. 2. design of the study

Only one independent factor is manipulated, with two modalities: some patients are in the VR group while the other patients are in the Non-VR group.

C. VR Used and Protocol

The VR device used for patients in the VR group is a stand-alone virtual reality headset from Lumeen® CE marked medical device.

Before debonding, each practitioner is instructed about the study objectives. All the patients will be debonded by a constant team (five residents and three seniors). They must remove the appliances with common methods. If there is no contraindication, VR system will be proposed to the patient by the practitioners before the orthodontic debonding. As Figures 3 and 4 show, patients will have the choice among nine immersive universes (five for adults relaxation and four dedicated to pediatric care). All were developed in cooperation with hypnotherapists and anesthesiologists. The presence of medical contraindications and refusal by a patient will be registered. the VR headset will be set up by the practitioner after installation on the dental chair and removed at the end of the procedure. The immersion will be stopped anytime for medical reasons or by simple request from the patient.

Fig. 3. immersive relaxing environments dedicated to pediatric care

Fig. 4. immersive relaxing environments developed for adults

D. Assessment of Pain and Dependent Factors

Before debonding, each participant is asked to complete (i) a Likert scale to assess the Anxiety score, as Figure 5 illustrates, (ii) the Numerical and Visual Rating Scale to assess pain scores, as Figure 6 shows. After the debonding procedure, patients will complete a more precise score for each dental area, as Figure 7 shows. So, there are three dependent factors.

Fig. 5. Visual Analog Scale (VAS) and Numerical Rating Scale (NRS), used to assess patients anxiety during debonding (usable for both adults and children).

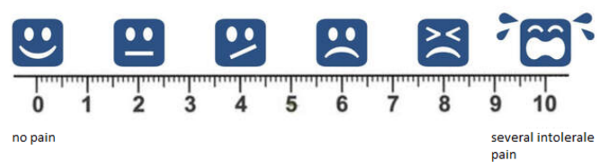

Fig. 6. Visual Analog Scale (VAS) and Numerical Rating Scale (NRS), used to assess patients pain during debonding (usable for both adults and children).

All the patients will be debonded by a constant medical team. Practitioners will remove the appliances with common methods, without adjunctive measures (finger pressure, biting cotton roll, music or video).

IV. Conclusion and Perspectives

Several studies have already shown that VR distraction can decrease patients' pain perception and anxiety. However, because there are very few studies dedicated to the effectiveness of VR on pain management during orthodontic debonding, the purpose of our study is to compare the patient reported pain and anxiety during their orthodontic appliance debonding with or without the use of a virtual reality headset.

For the moment (January 2024), we have included in the ”No-VR” group 15 patients with metal brackets, and 11 patients with ceramic ones.

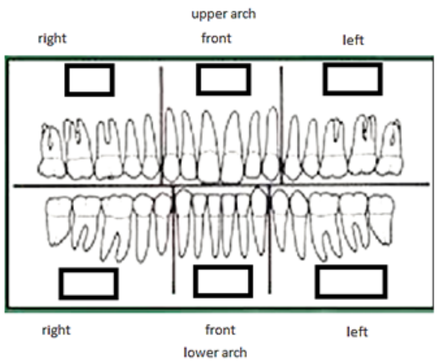

Fig. 7. Oral cavity diagram used to assess patients' pain through Numerical Rating scale (NRS) within each sextants.

To our knowledge, our experimental study is the first one investigating the effectiveness of the VR headset during orthodontic debonding. Of course, the potential impact of our research is not limited to orthodontic and dentistry patients. We hope that our future results will encourage speculation that VR pain reduction will also generalize to other acute pain populations/etiologies (e.g., oral and other ambulatory surgery, chemotherapy, radiation therapy) [42].

Acknowledgment

Thanks are due to Claire Coutureau and Baptiste Duceau for valuable discussions.

References

[1] P. Prasad P. M. S. Pradhan S. P. Gupta, S. Rauniyar. ”a randomized controlled trial to evaluate the effectiveness of different methods on pain management during orthodontic debonding”. *Progress in Orthodontics*, vol. 23(1):pp. 1–10, 2022.

[2] D. C. Turk E. Fernandez. ”the utility of cognitive coping strategies for altering pain perception: a meta-analysis”. *Pain*, vol. 38(2):pp. 123–135, 1989.

[3] S. Shahrbanian et al. ”use of virtual reality (immersive vs. non immersive) for pain management in children and adults: A systematic review of evidence from randomized controlled trials”. *PainEuropean Journal of Experimental Biology*, vol. 2(5):pp. 1408–1422, 2012.

[4] J. Swidrak M. V. Sanchez-Vives T. Donegan, B. E. Ryan. ”immersive virtual reality for clinical pain: Considerations for effective therapy”. *Frontiers in Virtual Reality*, vol. 1(9), 2020.

[5] V. Kapa J. Beecher-S; R. Sharar H. G. Hoffman, A. Garcia-Palacios. ”immersive virtual reality for reducing experimental ischemic pain”. *International Journal of Human-Computer Interaction*, vol. 15(3):pp. 469–486, 2003.

[6] Z. Khurshid E. Imran, N. Adanir. ”significance of haptic and virtual reality simulation (vrs) in the dental education: a review of literature”. *Applied Sciences*, vol. 11(21):10196, 2021.

[7] H. Koolivand et al. ”comparison of the effectiveness of virtual reality-based education and conventional teaching methods in dental education: a systematic review”. *BMC Medical Education*, vol. 24(1):pp. 8, 2024.

[8] S. Veazey et al. ”the development and usability assessment of an augmented reality decision support system to address burn patient management”. *BioMedInformatics*, vol. 4(1):pp. 709–720, 2024.

[9] V. J. Chen J. I. Gold A. Li, Z. Montano. ”virtual reality and pain management: current trends and future directions”. *Pain management*, vol. 1(2):pp. 147–157, 2011.

[10] M. I. Karobari et al. ”comparative evaluation of different numerical pain scales used for pain estimation during debonding of orthodontic brackets”. *International Journal of Dentistry*, 2021:pp. 1–10, 2021.

[11] W. J. da Silva Ursi D. Normando T. S. Normando, F. S. Calcada. ”patients' report of discomfort and pain during debonding of orthodontic brackets : a comparative studiy of two methods”. *World journal of orthodontics*, vol. 11(4):pp. 29–34, 2010.

[12] A. M. Pringle, A. Petrie, S. J. Cunningham, and M. McKnight. ”prospective randomized clinical trial to compare pain levels associated with 2 orthodontic fixed bracket systems”. *American Journal of Orthodontics and Dentofacial Orthopedics: Official Publication of the American Association of Orthodontists, Its Constituent Societies, and the American Board of Orthodontics*, vol. 136(2):pp. 160–167, August 2009.

[13] A. Balabanovsky A. D. Vardimon T. Brosh, A. Kaufman. ”in vivo debonding strength and enamel damage in two orthodontic debonding methods”. *Journal of Biomechanics*, vol. 38(5):pp. 1107–1113, May 2005.

[14] V. Krishnan. ”orthodontic pain: from causes to management–a review”. *European Journal of Orthodontics*, vol. 29(2):pp. 170–179, April 2007.

[15] Selma Elekdag-Turk. ”pain and/or Discomfort During Debracketing: A Review”. *Turkish Journal of Orthodontics*, vol. 32(4):pp. 236–240, July 2019.

[16] O. L. Williams and S. E. Bishara. Patient discomfort levels at the time of debonding: a pilot study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 101(4):313–317, April 1992.

[17] Y. Shen G. Shang L. Fang R. Wang J. Wang, X. Tang and Y. Xu. ”the correlations between health-related quality of life changes and pain and anxiety in orthodontic patients in the initial stage of treatment”. *BioMed Research International*, page 725913, 2015.

[18] B. Bradford S. Wilson P. Ngan, B. Kess. ”perception of discomfort by patients undergoing orthodontic treatment”. *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 96(1):pp. 47–53, July 1989.

[19] D. F. Brown and R. G. Moerenhout. ”the pain experience and psychological adjustment to orthodontic treatment of preadolescents, adolescents, and adults”. *Am J Orthod Dentofacial Orthop*, vol. 100:pp. 349–356, October 1991.

[20] B. Dinçer A. M. E. Erdinç. ”perception of pain during orthodontic treatment with fixed appliances”. *European Journal of Orthodontics*, vol. 26(1):pp. 79–85, February 2004.

[21] N. Nakada et al. ”pain and removal force associated with bracket debonding: a clinical study”. *Journal of applied oral science*, vol. 29:e20200879, 2021.

[22] A. Ulhaq F. Alharbi S. Alomari H. Mohammed M. Almuzian, M. Z. Rizk. ”effectiveness of different debonding techniques and adjunctive methods on pain and discomfort perception during debonding fixed orthodontic appliances: a systematic review”. *European Journal of Orthodontics*, vol. 41(5):pp. 486–494, September 2019.

[23] B. Bagis Y. Ren H. B. Pont, M. Özcan. ”loss of surface enamel after bracket debonding: An in-vivo and ex-vivo evaluation”. *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 138(4):pp. 387.e1–387.e9, October 2010.

[24] G. Sayar D. D. Kilinç. ”evaluation of pain perception during orthodontic debonding of metallic brackets with four different techniques”. *Journal of Applied Oral Science*, Vol. 27:e20180003, January 2019.

[25] N. Sahoo. ”comparison of the Perception of Pain during Fixed Orthodontic Treatment with Metal and Ceramic Brackets”. *Journal of Pharmacy Bioallied Sciences*, vol. 11(1):pp. 30–35, February 2019.

[26] H. Goh B. S. Boyd B. Mallari, E. K. Spaeth. ”virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis”. *Journal of Pain Research*, vol. 12:pp. 2053–2085, July 2019.

[27] E. C. Cappiello W. R. Camann D. Varelmann, C. Pancaro. ”nocebo-induced hyperalgesia during local anesthetic injection”. *Anesthesia and Analgesia*, vol. 110(3):pp. 868–870, March 2010.

[28] F. Wang et al. ”negative words on surgical wards result in therapeutic failure of patient-controlled analgesia and further release of cortisol after abdominal surgeries”. *Minerva Anestesiologica*, vol. 74(7-8):pp. 353–365, 2008.

[29] C. Alaterre et al. ”virtual Reality for PEripheral Regional Anesthesia”. *Journal of Clinical Medicine*, vol. 9(1):pp. 215, January 2020.

[30] A. C. Williams R. Ulrich T. J. Somers S. Malenbaum, F. J. Keefe. ”pain in its environmental context: Implications for designing environments to enhance pain control”. *PAIN*, vol. 134(3):pp. 241–244, February 2008.

[31] R. Fuzier, M. Lavidale, B. Bataille, A.-S. Richez, and J.-P. Maguès. [Anxiety: an independent factor of axillary brachial plexus block failure?]. *Annales Francaises D’anesthesie Et De Reanimation*, 29(11):776–781, November 2010.

[32] H. Demirci et al. ”watching a movie or listening to music is effective in managing perioperative anxiety and pain: a randomised controlled trial”. *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 31(12):pp. 6069–6079, December 2023.

[33] D. A. Thomas J. I. Gold, K. A. Belmont. ”the neurobiology of virtual reality pain attenuation”. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, vol. 10(4):pp. 536–544, August 2007.

[34] H. G. Hoffman et al. ”virtual Reality as an Adjunctive Non-pharmacologic Analgesic for Acute Burn Pain During Medical Procedures”. *Annals of Behavioral Medicine*, vol. 41(2):pp. 183–191, April 2011.

[35] A. Muayyad M. Eslam Bani. ”virtual reality as a distraction technique for pain and anxiety among patients with breast cancer: A randomized control trial”. *Palliative & Supportive Care*, vol. 17(1):pp. 29–34, February 2019. Publisher: Cambridge University Press.

[36] E. Kokinsky K. Enskär S. Nilsson, B. Finnström. ”the use of Virtual Reality for needle-related procedural pain and distress in children and adolescents in a paediatric oncology unit”. *European Journal of Oncology Nursing*, vol. 13(2):pp. 102–109, April 2009. Publisher: Elsevier.

[37] R. D. Goldman K. Arane, A. Behboudi. ”virtual reality for pain and anxiety management in children”. *Canadian Family Physician Medecin De Famille Canadien*, vol. 63(12):pp. 932–934, December 2017.

[38] F. D. S Costa N. B. Custódio and M. L. Goettems M. G. Cademartori, V. P. P. da Costa. ”effectiveness of Virtual Reality Glasses as a Distraction for Children During Dental Care”. *Pediatric Dentistry*, vol. 42(2):pp. 93–102, March 2020.

[39] M. Espí Mayor M. Miegimolle Herrero A. Constantini Leopardi, A. Adanero Velasco. ”effectiveness of Virtual Reality Goggles as Distraction for Children in Dental Care—A Narrative Review”. *Applied Sciences*, vol. 13(3):1307, January 2023. Publisher: Multidisciplinary Digital Publishing Institute.

[40] K. Le Du H. Vanquaethem, A. L. Septans. ”impact de la réalité virtuelle thérapeutique dans la gestion de la douleur et de l’anxiété liées aux soins en hématologie : focus chez les patients âgés de 75 ans et plus”. *La Revue de Médecine Interne*, vol. 44:pp. 95–96, June 2023.

[41] N. Simmons T. A. MacKenzie G. Kakoulides K. Bekelis, D. Calnan. ”effect of an Immersive Preoperative Virtual Reality Experience on Patient Reported Outcomes: A Randomized Controlled Trial”. *Annals of Surgery*, vol. 265(6):pp. 1068–1073, June 2017.

[42] H. G. Hoffman et al. ”the effectiveness of virtual reality pain control with multiple treatments of longer durations: A case study”. *International Journal of Human-Computer Interaction*, vol. 13(1):pp. 1–12, 2001.

Basic Senses and Their Implications for Immersive Virtual Reality Design

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

Jérôme Dinet
Université de Lorraine
Nancy, France
Email: jerome.dinet@univ-lorraine.fr

Anissa Bystrzycki
Université de Lorraine
Nancy, France
Email: anissa.bystrzycki4@etu.univ-lorraine.fr

Anne-Marie Eby
Université de Lorraine
Nancy, France
Email: anne-marie.eby6@etu.univ-lorraine.fr

Marie Gérard
Université de Lorraine
Nancy, France
Email: marie.gerard1202@gmail.com

***Abstract*— Experiencing immersion in daily life, interacting with other agents, holds significant value. The object of the interaction may be an agent in the real world or one in the virtual world. Such experiences occur when sensory stimuli input from the five senses are integrated in the brain, and cognitive processing based on the results of that integration leads to the expression of selected behaviors, which are then accompanied by changes in the environment. A sense of immersion is created by the seamless integration of changes in the environment and in one's own state. In this study, we leverage the Model Human Processor with Realtime Constraints (MHP/RT), a cognitive architecture that can simulate human action selection and encompasses perceptual, cognitive, motor, and memory processes, and we argue that within the integration of multimodal sensory stimuli, *P-Resonance* – a resonance occurring between environmental information and perceptual memory, and *C-Resonance* – a subsequent resonance involving cognitive processes characterized by Two Minds and memory – are pivotal components contributing to the sense of immersion. The state of the external world is perceived as M-dimensional information by the five senses. This is passed to cognitive processing via P- and C-Resonances, converted into N-dimensional motor information in memory. Further, it is physically executed in the external world. We argue that the basic senses of rhythm, space, and number exist as mechanisms for these resonances to occur. It is shown that the spatial sense involving vision, not functioning before human birth, is integrated after birth with the rhythmic sense, which had been developing before birth. Furthermore, the number sense involving higher-order object recognition is further integrated with growth, which suggests that these basic senses are involved in the basis for the expression of human behavior. We contend that a virtual environment that does not deviate from the basic sensory-based mechanisms, which manifest interactions in the real world, is essential for the generation of immersive experiences.**

***Keywords*— Basic senses; Rhythmic sense; Spatial sense; Number sense; Resonance.**

I. Introduction

The environment in which humans interact includes 1) *natural objects* that are direct objects of interaction, 2) *real artifacts* created by humans that exist as entities, 3) *virtual artifacts* that do not exist as entities but are perceived by humans through the illusion of their existence, and 4) real objects and artifacts that exist naturally and are indirectly involved in interaction as their background. Of these four categories, the third, "virtual artifacts," is rapidly permeating our lives. This change is attributed to the advancement of artificial intelligence algorithms, which now enable the creation of intricate illusions.

An interface exists between humans and the environment. When the interface is conscious, or visible, the human consciously performs actions toward the environment and consciously evaluates the response of the environment to those actions. In this case, the interface is visible to the human. Conversely, there are cases in which interaction proceeds without the human being aware of the interface. In this case, the interface is transparent or invisible, the human is immersed in the environment. Moreover, the human and the environment are seamlessly integrated.

When two interacting systems are seamlessly integrated, the processes in each system proceed in unison and in step with each other. The two systems are then synchronized. When both systems are linear, synchronization is easy because the progress of each process can be controlled based on highly accurate predictions. However, if either system is nonlinear, synchronization is challenging. Notably, human behavior can be captured by four nonlinearly connected hierarchies or bands, that is, biological, cognitive, rational, and social bands [1]; thus, humans are nonlinear behavior-generating systems. In the case of action selection, it can be captured by Two Minds [2][3], comprising a conscious system and an unconscious system. The conscious system is called System 2. This is a slow system that operates by feedback control, which executes inference using knowledge, whereas the unconscious system is called System 1 and is a fast system that operates by feed-forward control using intuition.

The behaviors that humans produce in their interactions with the environment are the result of integrating the behaviors of System 1 and System 2, which have different characteristic times. Regardless of whether the system existing as an environment is a linear or nonlinear system, it is necessary to keep pace with human behavior to create a sense of immersion on the part of humans. This can be achieved by establishing a seamless relationship with the nonlinear human system.

When evaluating the alignment of the environment and humans states, it is not appropriate to simply consider that the actions of both systems are in sync with each other at time T. For an interaction event occurring at a certain time, humans consciously plan in advance and unconsciously coordinate and execute their actions just before the interaction. They also unconsciously make adjustments in the neural networks related to the generation of the action after the interaction. Further, they consciously reflect on the interaction when the outcome of the interaction is obtained. The conscious and unconscious learning after the interaction is used to plan actions when similar situations are encountered in the future [4]. From the human perspective, this event occurs within the time range of the Two Minds process operating before and after the event. Therefore, it is reasonable to view the synchronization between the environment and humans involved in the event generation within that time range. This type of synchronization with a time width is called *weak synchronization*. Further, it is proposed as a necessary condition for the generation of a sense of immersion. *Immersive feeling eliciting condition* for an artificial environment to have the user feel immersive-ness is defined as follows [5]:

IMMERSIVE FEELING ELICITING CONDITION

1) It must be new to them.
2) **With an anticipation activated by the artificial environment, they are able to perform actions without any breakdown in performing motor-level actions.**
3) They are able to consciously recognize an event associated with the series of just-finished actions.
4) They are able to reflect on the event to integrate it with the recognized feeling associated with the event.

The longer the state shown in Item 2 of the immersive feeling eliciting condition, the longer the immersive sensations can be felt without conscious intervention by System 2. Here, a series of actions is generated by feedforward control by System 1 as a series that is executed unconsciously without evaluation of the execution results. Let us assume that the first action is executed at time T and the last action in this series is executed at time T'. Subsequently, the interaction event generated by this action series is denoted by $E(T : T')$. The period $[T,\ T']$ when the action is executed unconsciously is the time when the action does not come to consciousness. For instance, the time when the action is interacting with the environment under feedforward control by System 1 without conscious intervention by System 2, and it can be regarded as the immersion time. The longer this time is, the longer the immersion is held.

The conditions for the establishment of weak synchronization that enables a continuous sense of immersion in the interaction with virtual artifacts depend on the perceptual, cognitive, and motor characteristics of the individual, including the content of memory, and the individual characteristics of the combination of these characteristics. This study examines how immersion is continuously maintained assuming general individual characteristics, based on Model Human Processor with Realtime Constraints (MHP/RT) [4][6]. MHP/RT is a comprehensive theory of action selection and memory, and provides the basis for building any model to understand human everyday behavior, including cognitive mechanism. Then, based on this, we discuss some individual characteristics that are unique but should be appropriately supported, and examine means to generate a sense of immersion.

This paper is organized as follows: Section II describes the outline of information uptake by perceptual processes from the external and internal environment, memory activation and execution of cognitive and motor processes through resonance in Section II-A. Section II-B introduces an important problem of binding and integrating multiple sensory information to produce behavior synchronized with spatiotemporal changes in the environment. Section II-C argues that the basic senses, that is, the rhythmic, spatial, and number senses, are critical for incorporating external information through resonance to make it available for the cognitive and motor processes to follow. Section III discusses the basic senses from the perspective of human development, including the prenatal period. Section IV concludes this paper by suggesting the points to be considered when designing virtual reality that smoothly connects to human activities.

II. INTERACTING WITH THE ENVIRONMENT USING RESONANCE

Satisfying the second item of Immersive Feeling Eliciting Condition (IFEC) is important for seamless interactions to continue to run as a feedforward process by System 1. The key for understanding the human–environment interaction process based on MHP/RT is that the communication between autonomous systems is achieved by a mechanism of *resonance* [7]. Both environmental systems and human systems are autonomous systems; human systems include perception, cognition, movement, and memory. This section describes how the continuous feedforward processing behind seamless interactions is supported by resonance mechanisms.

A. Interaction with the Environment Through Memory, Perception, Cognition, and Motor Processes Using Resonance

When interacting with the environment, humans respond to physical and chemical stimuli emitted from the external and internal environment by sensory nerves located at the interface with the environment and take in environmental information in the body. The brain acquires environmental information concerning the current activity of the self through the multiple

 ISBN: 978-1-68558-154-1

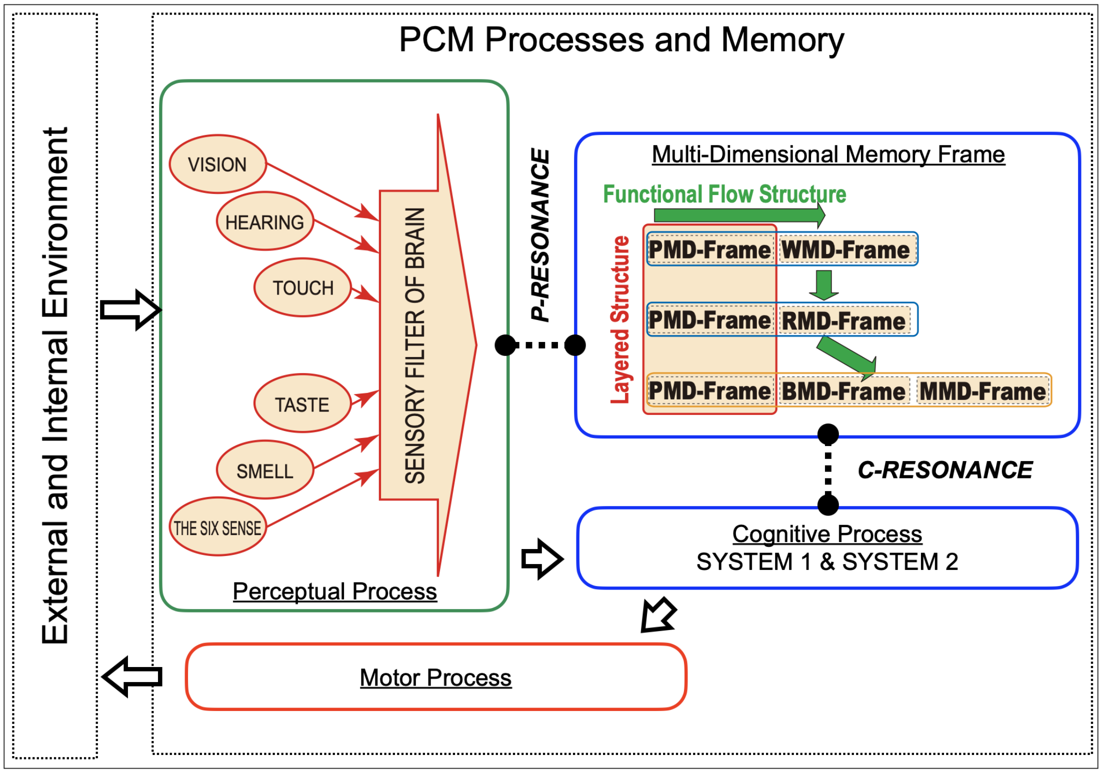

Figure 1. Information uptake by perceptual processes from the external and internal environment, memory activation and execution of cognitive and motor processes through resonance.

sensory organs. 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 shows the process, based on MHP/RT [6][8], by which environmental information, internal and external, is taken into the body via sensory nerves, processed in the brain, and then acted upon by the external world via motor nerves. This process involves memory, which is modeled as Multi-Dimensional Memory Frame, and perceptual, cognitive (Two Minds), and motor processes. The memory structure, Multi-Dimensional Memory Frame, consists of Perceptual-, Behavior-, Motor-, Relation-, and Word-Multi-Dimensional Memory Frame. Perceptual-Multi-Dimensional Memory Frame overlaps with Behavior-, Relation-, and Word-Multi-Dimensional Memory Frame, for spreading activation from Perceptual-Multi-Dimensional Memory Frame to Motor-Multi-Dimensional Memory Frame.

Perceptual information taken in from the environment through sensory organs resonates with information in the memory network structured as Multi-Dimensional Memory Frame, that is, P-Resonance. In Figure 1, this process is indicated by •—•. Resonance occurs first in the Perceptual-Multi-Dimensional Memory Frame within the Multi-Dimensional Memory Frame and activates the memory network. After that, the activity propagates to the memory networks that overlap the Perceptual-Multi-Dimensional Memory Frame, which are the Behavior-, Relation-, and Word-Multi-Dimensional Memory Frame, and finally to the Motor-Multi-Dimensional Memory Frame. In cognitive processing by Two Minds, conscious processing by System 2, which utilizes the Word- and Relation-Multi-Dimensional Memory Frame via C-Resonance, and unconscious processing by System 1, utilizing the Behavior- and Motor-Multi-Dimensional Memory Frame via C-Resonance, proceed in an interrelated manner. The motor sequences are expressed according to the Motor-Multi-Dimensional Memory Frame, which is the result of cognitive processing. The memories involved in the production of a behavior are updated to reflect the traces of its use process and influence the future behavior selection process.

B. Integration of Multiple Sensory Information to Produce Behavior Synchronized with Spatiotemporal Changes in the Environment

Information from the environment is taken in by multiple sensory organs. The sensory organs are located in different parts of the body with a certain extent, and the information collected is spatially distributed. Additionally, the information received by the sensory organs is time-series information. Therefore, the information taken in through the sensory organs is characterized as information that is spatially and temporally spread out. *How, then, do the structured Multi-Dimensional*

Memory Frame and perceptual information resonate? Or how can P-Resonance occur?

This question is known as the binding problem. It is a question of following form: *How does the brain integrate different aspects of perception into a coherent experience?* or *How does it perceive an object as a single entity while having distinguishable aspectual features such as shape, color, and texture?*

1) Dennett's Thoughts on the Binding Problem: Related to this question, Daniel Dennett, author of "Consciousness Revealed" [9], was asked by an interviewer, "You've talked about consciousness as something that we feel that's really more marvelous than it really is. Why is it?" To this, he responded as follows:

> *Yes, I think that consciousness plays tricks on us. Behind the eyes and in between the ears and the inner witnesses watching this wonderful show. But then when you do the, the physiology and you study perception, you realize it now in fact, you have a very limited take, you're only taking sips from them firehose of information that's coming in a little bit from vision a little bit from from hearing, and there's in fact this, this competition going on tug of war between different senses, between different interpretations of what you're seeing. All of this competition resolves itself in the fullness of time and pretty darn quick to produce the behavior that we're capable of and the reflection that we're capable of. It seems though, as if there's, it all comes together at some place.* ***And that's just an illusion. There's no place in the brain where it all comes together, for enjoyment and for and for witnessing by an inner witness.*** *So we have to take all that work that that inner witness was going to do, and we have to break it up in little bits and distributed around in the brain in the time and space that's available. And no one of those little bits is going to be conscious, and yet the sum of all of that work. Witnesses are going to do that with consciousness.*

Subsequently, he adds the following to the "illusion that everything flows together as one":

> *We have that illusion, and it is an illusion, because in fact, there's lots of things going on at once. They're not all that coherent.* ***And so our brains are very good at creating the sorts of simplifications that make that make its own job easier.*** *And so, yes, consciousness is an illusion of the brain for the brain, by the brain, if you like.*

2) Mechanistic Answers to the Binding Problem: Our answer to the binding problem is "sensory filter processing of multiple perceptual information and memory activation via *P-Resonance*." In the theorization of human perceptual, cognitive, and motor processes and memory by MHP/RT, the interface between the environment and the brain is performed by P-Resonance in perceptual information processing. Moreover, the interface between hierarchies in non-linearly connected hierarchical information processing within the brain is performed by C-Resonance. Resonance expresses the relation between two sides connected in a nonlinear relationship. Furthermore, it is a mechanism to transfer information that is ordered on one side to be treated as ordered information on the other side by mapping information. Notably, each information spreads with a specific manner on a common time axis [10].

C. Theoretical Derivation of Basic Senses

MHP/RT can be thought of as implementing the method described by Daniel Dennett in Section II-B1 on a system consisting of perceptual, cognitive, and motor processes and memory. Given the characteristics of memory described in Section II-B2, this section discusses perceptual memory, which is activated by P-Resonance. Further, it shows that *basic senses* exist to process perceptual information from sensory organs in an orderly manner. Basic senses include *rhythmic sense* related to time, to be described in II-C6a, *spatial sense* related to spatial perception, to be described in II-C6b, and *number sense* related to object segmentation in space, to be described in II-C6c.

1) Information that Perception Takes in: There are two types of information that perception takes in, which is represented as a discrete sequence of points spread over a time axis:

1) Stationary periodic sampling data; and
2) Differential information when it senses changes in the environment.

Moreover, there are two types of perception that involve different sources of information:

1) Perception for environmental information outside the body; and
2) Perception for monitoring the state of activity inside the body.

Therefore, the objects to be perceived can be broadly classified into four types depending on whether the data is stationary or non-stationary, and whether the source is inside or outside the body. There are various perceptual organs; each collecting its own perceptual information within its own perceptible bandwidth.

2) Memory Generation from Perceived Information: The perceived information concerns generation of memory for bodily behavior. Initially, memories are formed for different organs, and even within a single cell, in different forms. Nevertheless, human bodily behavior in general can be understood almost entirely through the workings of the central nervous system, which comprises the brain and spinal cord.

However, it is known from genetic analysis that the underlying structure of the nervous system is a chain of sensory, intervening, and motor nerves. An important function of the perception originated from sensory nerves is the identification of external objects to be transmitted to intervening nerves. The information from various perceptions is integrated to confirm the existence of the object and to store it in memory. Therefore, most memories that reside in the cerebral cortex in brain are object memories.

3) $M \otimes N$ Mapping in Memory: The memory formed by the chain structure of sensory, intervening, and motor nerves stores the procedures of actions performed in response to events in the intervening stage, situated between perception and motor movement, by means of pairwise relations between perceptual components of the objects expressed as the M-dimensional information, that is, *perceptual objects*, and motor components of the objects expressed as the N-dimensional information, that is, *motor objects*. These procedures expressed in the intervening nerves can be viewed as chains in the Behavior-Multi-Dimensional Memory Frame. These are collectively referred to as "$M \otimes N$ mapping in memory" in MHP/RT [11].

4) Two Minds: System 1 and System 2: In the range of everyday behavior, processing is known to occur in the basal ganglia, that is, a group of neuronal nuclei that connect the cerebral cortex to the thalamus and brain stem, and below, where the $M \otimes N$ mapping in memory is coordinated by Two Minds. System 1 and System 2 work in parallel. The System 1 process is mainly responsible for physical activities, whereas the System 2 process works as an intervention for the purpose of coordinating activities. The role of System 2 is to intervene effectively in physical actions. This is accomplished by enhancing perceptual abilities through reconstruction of perceptual memory organized as Perceptual-Multi-Dimensional Memory Frame. For this purpose, System 2 consciously reflects on the results of physical actions conducted for accomplishing then-activated goals and reconstructs Perceptual-Multi-Dimensional Memory Frame according to the evaluation.

5) Representing Perceptual and Motor Objects in Relativized Time and Circulation Networks: The perceptual information taken in by sensory organs *does not contain absolute positional and temporal information*. The only shared feature among the various types of perceptual information is *simultaneity with other perceptual information* in the same time axis. Specifically, parallelism and cross-synchronization between each perceptual information processing are ensured. Additionally, it is possible to execute actions coherently as a procedure by integrating various types of perceptual information in a certain period of time on the time axis.

Human behavior forms a cyclic life ecology called circadian rhythm with a fluctuating bandwidth, which is caused by the stable day–night periodicity of the earth's rotation, the inclination of the axis of rotation with respect to the sun's orbital plane, and the slightly elliptical movement of the orbital circle. Human behavior contains recursive elements in which the results of one's actions are returned to oneself. Therefore, System 1 can act adaptively and flexibly in the next similar situations, whereas System 2 can obtain new effective action procedures by reflecting on the results of the action and organizing the perceptual information accordingly to reconstruct the existing memory network, that is, the Perceptual-Multi-Dimensional Memory Frame. In the future, System 2 intervenes in the sequence of actions generated by System 1 when it is needed to change its direction based on the reconstructed memory.

Human behavior can be viewed as an adaptive behavior for survival in the ever-changing and pseudo-cyclic environment. From birth to death, human beings act ceaselessly, which is a cyclic activity encompassing circadian rhythms. Hence, the neural network concerned with the execution of System 1 forms a circulatory network. Humans transform their behavior as their environment changes and as they themselves change with age. Accordingly, the initial circulatory network develops into a circulatory network with more complex connections and connection bandwidths by connecting the neural circuits involved in the execution of System 2 as procedures. Conscious procedure performed by System 2 happens in the operation and unit task systems in the cognitive band, and rational and social bands [1]. In MHP/RT, such procedures are represented in the Relation-Multi-Dimensional Memory Frame as consciously accessible relations of manipulable objects.

6) Memory Reuse: Rhythm-Based Reconstruction of Cognitive Objects: The earth is almost spherical. However, the range of human activity is far smaller than the global scale. Therefore, it is perceived as a three-dimensional Cartesian space rather than a coordinate system associated with a sphere.

Memories do not contain absolute temporal and three-dimensional positional information as described in Section II-C5. Meanwhile, System 1 and System 2 have to generate timely behavior appropriate for the time-position dependent situation. It is necessary to reconstruct the information from the memories by incorporating the time and positional information to make the time-position free memories available for System 1 and System 2 in the time position dependent ongoing situation. By doing this, System 1 and System 2 can reuse the time-position free memories at the point in time when it becomes necessary.

In Figure 1, the Multi-Dimensional Memory Frame and System 1 and System 2 are connected via C-Resonance, that is, the cognitive process incorporates the portion of memory through resonance that has been transformed into manipulable forms in the current situation by binding the time and position information in P-Resonance, which is called *Cognitive-Objects* in MHP/RT [12]. It is considered that, whereas perception of the external situation resonates with the Perceptual-Multi-Dimensional Memory Frame via P-Resonance, *the time-position free constructs* in the Perceptual-Multi-Dimensional Memory Frame are somehow modified to generate *the time-position dependent constructs*, that is, Cognitive-Objects, that are necessary to instantiate the real actions stored in the Motor-Multi-Dimensional Memory Frame via Relation- and Behavior-Multi-Dimensional Memory Frame. Then, how can it be possible? The key is the concept of *rhythm*, which characterizes the timing of the occurrence of an event. That is to say "binding positional and time information to the time-position free information by means of rhythm."

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.

b) Spatial Sense through Rhythmic 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, that can be called "spatial 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.

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 called number sense [13]. Perceptual information is represented in M-dimensional information in the Perceptual-Multi-Dimensional Memory Frame via P-Resonance with 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 [5]. Subsequently, they are aggregated as cognitive objects by the number sense and made available to Two Minds via C-Resonance.

d) Summary: 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.

III. THE ROLE OF RESONANCE IN THE HUMAN DEVELOPMENT

Following birth, infants must immediately process and rapidly adapt to the array of unknown sensory experiences associated with their new ex-utero environment. However, and as Dall'Orso and colleagues (2020) [14] said, although it is known that unimodal stimuli induce activity in the corresponding primary sensory cortices of the newborn brain, it is unclear how multimodal stimuli are processed and integrated across modalities.

Despite the relatively immature state of their sensory modalities, newborns have perceptions that are acquired, and are triggered by, their contact with the environment. More recently, the study of the fetal origins of the sensory modes has revealed that in utero all the senses prepare to operate, except for the vision mode, which is only functional starting from the first minutes after birth [15][16]. This discrepancy between the maturation of the different senses leads to the question of how human newborns come to understand our multimodal and complex environment.

According to us, and as we wrote previously in this paper, the information taken from the sensory organs is characterized as information that is spatially and temporally spread out and we hypothesize that the Multi-Dimensional Memory Frame and perceptual information resonate (P-resonance) can be relevant to describe how multimodal stimuli are processes and integrated.

Just after the birth, the human newborn nervous system must also quickly adapt to process the temporal, spatial, and contextual features of complex inputs which simultaneously stimulate different sensory systems. This integration between different sensory modalities is thought to be established through associative learning during postnatal life as it allows simple and fast encoding of environmental contingencies [17].

If this integration between different sensory modalities has multiple roles in early development for biological and physiological aspects, for example, through anticipating and overcoming respiratory or thermal challenges during sleep, it plays also a key role for cognitive and social development such as behavioral and emotional-self-regulation [18], attention to facial and vocal expressions [19], linking different auditory–visual features to make inferences about specific objects [20][21], and facilitating development of language and vocabulary growth [22].

Dall'Orso and colleagues (2020) [14] have demonstrated that the early human brain is already capable of processing external and simultaneous multisensory information within the two distinct primary sensory cortices. We further found that this multisensory stimulus presentation can directly influence cortical activity in a crossmodal manner, even in the absence of its archetypal substrate. They also found that the process of encoding information during associative learning in the neonatal period engages wider cortical association regions across relevant neural networks. Recent Magnetoencephalography (MEG)/Electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) studies have suggested that while posterior parietal cortices integrate signals weighted by their sensory reliabilities irrespective of task context, anterior parietal cortices encode spatial estimates depending on their reliability and task relevance [23][24][25]. In anterior parietal cortices, spatial estimates rely more on the location of the signals of the sensory modality that needs to be reported. While these behavioral and neuroimaging findings demonstrate that observers' perceptual goals influence how the brain combines sensory signals to support perceptual inference, the underlying computational and neural mechanisms remain controversial and we argue that MHP/RT can provide a relevant theoretical framework to explain these mechanisms.

And finally, their results confirmed prior results obtained by other authors who identified the primary sensorimotor cortices as the encoders of our associative learning task, supporting the idea that they are not only involved in basic sensory perception but also higher level processes such as multisensory integration and experience specific memory encoding [26]. Moreover, some recent studies investigating multimodal integration for newborns with perceptual impairments, such as visual impairments [27][28][29] and for children with Autism Spectrum Disorder, or ASD [30], provide very similar results.

IV. Conclusion and Future Work

This study argued in Section II-B that binding position and time to the Multi-Dimensional Memory Frame is essential to generate situation-adaptive behavior, and derived in Section II-C basic senses for solving the binding problem. 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. Furthermore, 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.

Section III discussed the basic senses from a developmental perspective, tracing it back to the prenatal stage. It also clarified how we are able to process multimodal information now that we are living surrounded by various artifacts, which implements immersive virtual reality. Based on the argument this paper provided, establishing P-Resonance between the external environment and the Perceptual-Multi-Dimensional Memory Frame with the focus of the basic senses. These include the rhythmic sense, spatial sense, and number sense, should be critical for smooth integration of the environment and human being in action. As long as the source of stimuli to human sensory organs, whether virtual or real, is not smoothly processed by the basic senses, it will not be taken into the organism and processed in relation to it. If those stimuli are Immersive Virtual Reality generated by applying Artificial Intelligence, they must be generated to achieve some goal, but if they are not compatible with the basic senses, that goal cannot be achieved.

In everyday life human beings experience a continuous stream of information that they perceive through sight, sound, smell, taste, and touch. Even though this experience is mostly multisensory, that is, they receive information from multiple senses simultaneously, psychological research has primarily focused on studying our senses in isolation. Multisensory processing refers to the interaction of signals arriving nearly simultaneously from different sensory modalities. This implies that information from one modality can influence information processing in another modality. Information from different sensory modalities can also be combined into a single multisensory event, a process that is referred to as multisensory integration [31]. By this way, IVR is used to create virtual worlds that are as immersive as possible in order to make users feel as if they are "really there", immersion referring to the objective capacity of the technology to deliver sensorial stimulations and movement tracking, for example, head, hands, comparable to their physiological manifestations in the physical world [32]. It is assumed that the more immersed someone feels while exposed to a virtual world, then the more memorable the IVR experience is specially for children [33]. Even if some studies show that too much visual stimulations, for example, by a larger field of view or more visual details, can be responsible for negative effects and discomfort due to increased eye strain on the part of the user, that is, cybersickness, the quality of immersion results from the combination of various factors, such as field of view (the extent of the visible world), head tracking, visual fidelity (the realism and details of visual information) and multisensory information, which is thought to be a central element in immersion.

At the moment of sampling continuous events in space–time, which is the moment the memory is formed, absolute time and coordinate information is lost and associations between events are made with features that can be mapped to other events. Therefore, when the events are drawn out at the time of motion, spatio-temporal information between events is necessary, so it is interpolated to generate it. P-Resonance occurs between the external stimulus and the Perceptual-Multi-Dimensional Memory Frame by rhythmic and spatial senses, and M-dimensional perceptual representations are generated. Subsequently, under the condition of weak-synchronization, cognitive objects are generated by utilizing the number sense. C-Resonance occurs between the cognitive object and the Two Minds, and finally, mapped to the N-dimensional representation of the Motor-Multi-Dimensional Memory Frame. The N-dimensional motion representation is interpolated in space–time and converted into motor-enabled information to generate motion via motor nerves. The body plan, that is, skeleton, supports the whole behind the scenes as the base of the interpolation program as the default value. For memories created in a real environment to be effective in a VR environment, which would be IVR, VR design should be based on these characteristics.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 20H04290. The author would like to thank Editage (www.editage.com) for English language editing.

References

[1] A. Newell, Unified Theories of Cognition (The William James Lectures, 1987). Cambridge, MA: Harvard University Press, 1990.
[2] D. Kahneman, “A perspective on judgment and choice,” American Psychologist, vol. 58, no. 9, 2003, pp. 697–720.
[3] D. Kahneman, Thinking, Fast and Slow. New York, NY: Farrar, Straus and Giroux, 2011.
[4] M. Kitajima and M. Toyota, “Four Processing Modes of *in situ* Human Behavior,” in Biologically Inspired Cognitive Architectures 2011 - Proceedings of the Second Annual Meeting of the BICA Society, A. V. Samsonovich and K. R. Jóhannsdóttir, Eds. Amsterdam, The Netherlands: IOS Press, 2011, pp. 194–199.
[5] 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. New York, NY, USA: ACM, 2018, pp. 18/04:1–18/04:8. [Online]. Available: https://doi.org/10.1145/3234253.3234301
[6] M. Kitajima, Memory and Action Selection in Human-Machine Interaction. Wiley-ISTE, 2016.
[7] J. Dinet and M. Kitajima, “The Concept of Resonance: From Physics to Cognitive Psychology,” in COGNITIVE 2020 : The Twelfth International Conference on Advanced Cognitive Technologies and Applications, 2020, pp. 62–67.
[8] 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, 2013, pp. 82–93.
[9] D. C. Dennett, Consciousness Explained (English Edition). Little, Brown and Company, 2 2017.
[10] M. Kitajima et al., “Language and Image in Behavioral Ecology,” in COGNITIVE 2022 : The Fourteenth International Conference on Advanced Cognitive Technologies and Applications, 2022, pp. 1–10.
[11] 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, 2021, pp. 148–161.
[12] 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, 2012, pp. 41–58.
[13] S. Dehaene, The Number Sense: How the Mind Creates Mathematics, Revised and Updated Edition (English Edition). Oxford University Press, 4 2011.
[14] S. Dall’Orso et al., “Cortical Processing of Multimodal Sensory Learning in Human Neonates,” Cerebral Cortex, vol. 31, no. 3, 11 2020, pp. 1827–1836. [Online]. Available: https://doi.org/10.1093/cercor/bhaa340[retrieved:January,2023]
[15] A. Streri, M. D. d. De Hevia, V. Izard, and A. Coubart, “What do we know about neonatal cognition?” Behavioral Sciences, vol. 3, no. 1, 2013, pp. 154–169. [Online]. Available: https://www.mdpi.com/2076-328X/3/1/154[retrieved:January,2023]
[16] A. Streri and M. D. de Hevia, “How do human newborns come to understand the multimodal environment?” Psychonomic Bulletin & Review, vol. 30, no. 4, Aug. 2023, pp. 1171–1186. [Online]. Available: https://doi.org/10.3758/s13423-023-02260-y[retrieved:January,2023]
[17] K. Connolly, “Multisensory perception as an associative learning process,” Frontiers in Psychology, vol. 5, 2014. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01095[retrieved:January,2023]
[18] L. P. Lipsitt, “Learning and Emotion in Infants,” Pediatrics, vol. 102, no. Supplement_E1, 11 1998, pp. 1262–1267. [Online]. Available: https://doi.org/10.1542/peds.102.SE1.1262[retrieved:January,2023]
[19] A. S. Walker-Andrews, “Infants’ perception of expressive behaviors: Differentiation of multimodal information.” Psychological Bulletin, vol. 121, no. 3, 1997, pp. 437–456, place: US Publisher: American Psychological Association.
[20] B. A. Morrongiello, K. D. Fenwick, and G. Chance, “Crossmodal learning in newborn infants: Inferences about properties of auditory-visual events,” Infant Behavior and Development, vol. 21, no. 4, 1998, pp. 543–553. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0163638398900285[retrieved:January,2023]
[21] Y. Tanaka, Y. Kanakogi, M. Kawasaki, and M. Myowa, “The integration of audio − tactile information is modulated by multimodal social interaction with physical contact in infancy,” Developmental Cognitive Neuroscience, vol. 30, 2018, pp. 31–40. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1878929316301232[retrieved:January,2023]
[22] L. Smith and C. Yu, “Infants rapidly learn word-referent mappings via cross-situational statistics,” Cognition, vol. 106, no. 3, 2008, pp. 1558–1568. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010027707001795[retrieved:January,2023]
[23] T. Rohe and U. Noppeney, “Cortical hierarchies perform bayesian causal inference in multisensory perception,” PLoS biology, vol. 13, no. 2, 2015, p. e1002073.
[24] T. Rohe and U. Noppeney, “Distinct computational principles govern multisensory integration in primary sensory and association cortices,” Current Biology, vol. 26, no. 4, 2016, pp. 509–514.
[25] M. Aller and U. Noppeney, “To integrate or not to integrate: Temporal dynamics of hierarchical bayesian causal inference,” PLoS biology, vol. 17, no. 4, 2019, p. e3000210.
[26] D. B. Headley and N. M. Weinberger, “Relational Associative Learning Induces Cross-Modal Plasticity in Early Visual Cortex,” Cerebral Cortex, vol. 25, no. 5, 11 2013, pp. 1306–1318. [Online]. Available: https://doi.org/10.1093/cercor/bht325[retrieved:January,2023]
[27] N. J. Dale et al., “Home-based early intervention in infants and young children with visual impairment using the developmental journal: longitudinal cohort study,” Developmental Medicine & Child Neurology, vol. 61, no. 6, 2019, pp. 697–709. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/dmcn.14081[retrieved:January,2023]
[28] F. Morelli et al., “A multidimensional, multisensory and comprehensive rehabilitation intervention to improve spatial functioning in the visually impaired child: A community case study,” Frontiers in Neuroscience, vol. 14, 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2020.00768[retrieved:January,2023]
[29] A. Gatecel and A. Valentin-Lefranc, Le Grand Livre des pratiques psychomotrices. Fondements, domaines d’application, formation et recherche, ser. Hors collection. Paris: Dunod, 2019. [Online]. Available: https://www.cairn.info/le-grand-livre-des-pratiques-psychomotrices--9782100798940.htm[retrieved:January,2023]
[30] N. M. Salowitz, “Multimodal sensory integration for perception and action in high functioning children with autism spectrum disorder,” Ph.D. dissertation, Marquette University, 2015.
[31] M. Quak, R. E. London, and D. Talsma, “A multisensory perspective of working memory,” Frontiers in human neuroscience, vol. 9, 2015, p. 197.
[32] L. B. Cadet and H. Chainay, “Memory of virtual experiences: Role of immersion, emotion and sense of presence,” International Journal of Human-Computer Studies, vol. 144, 2020, p. 102506.
[33] L. B. Cadet, E. Reynaud, and H. Chainay, “Memory for a virtual reality experience in children and adults according to image quality, emotion, and sense of presence,” Virtual Reality, vol. 26, no. 1, 2022, pp. 55–75.

Model-Based Analysis of the Differences in Sensory Perception between Real and Virtual Space : Toward "Adaptive Virtual Reality"

Taichi Nakagawa
Nagaoka University of Technology
Nagaoka, Niigata, Japan
Email: s223308@stn.nagaokaut.ac.jp

Muneo Kitajima
Nagaoka University of Technology
Nagaoka, Niigata, Japan
Email: mkitajima@kjs.nagaokaut.ac.jp

Katsuko T. Nakahira
Nagaoka University of Technology
Nagaoka, Niigata, Japan
Email: katsuko@vos.nagaokaut.ac.jp

***Abstract*— Although the implementation of "Adaptive Virtual Reality" is becoming feasible, understanding the main effects of its realization on users based on cognitive models is essential. Here, as the first step, we first describe a model of the flow of information obtained by actual human perception through avatars in virtual reality (VR) and the resulting human reactions, and confirm the validity of the user models proposed so far. We also consider the degree of immersion predicted due to the integration of multimodal information. The cognitive processes of VR experiences are largely categorized into "perception and recognition of information (attention, memory, and decision making)" and "perception-based physical actions and interactions with VR objects". Based on this, we describe a cognitive model of VR experiences. In addition, as examples of the discrepancies in sensory perception experienced in real/VR spaces, we briefly describe the phenomena that occur in communication. We describe the cognitive models for these phenomena and qualitatively consider the degree to which sensory information obtained from the real/VR space affects the degree of chunks activation. The intensity of human sense is expressed as a logarithm according to Weber-Fechner's Law, suggesting that human senses can distinguish differences even with weak sensory information. We argue that the "slightly different from the real world" sense felt in VR content is caused by such slight differences in sensory information. Overall, we advance the cognitive understanding of the immersive experience particularly in the VR space, and qualitatively describe the possibility of designing highly immersive VR content which are adapted to each individual.**

Keywords*— *sensory perception; cognitive model; virtual reality; experience.

I. Introduction

With the growth in Virtual Reality (VR) goggles and low cost of equipment for shooting omnidirectional video, VR content has attracted substantial attention. In addition to games, a wide range of VR contents have been developed, including omnidirectional video playback, education, sightseeing, property previews, and shopping. VR systems that enable these contents to be viewed are also growing rapidly. For example, the following innovations have emerged in content design. VR systems using Head-Mounted Displays (HMDs) sold to general consumers cover the user's field of vision; thus, the user cannot see their own body. Therefore, VR systems using HMDs typically display a virtual body drawn from the user's first-person perspective. A mechanism for realizing the user's first-person perspective is the implementation of avatars. The effects of avatars have been described by researchers. Steed et al. [1] suggested that the use of avatars that follow the user's movements can reduce the cognitive load of certain tasks in the VR space. People around the world have been using VR social networking services, such as VRChat, where users enjoy interacting with other users using avatars that they have selected and edited to their liking. Theis shows that avatars are a means of self-expression in VR communication.

There are many research approaches to VR contents and systems, including research from the perspective of Human Computer Interaction (HCI), research on the differences in sensory perception between the real world and VR, and research on "Adaptive VR" that incorporates individual adaptability into VR contents.

Among the studies from the perspective of HCI, Mousavi et al. [2] integrate Emotion Recognition (ER) and VR to provide an immersive and flexible environment in VR. This integration can advance HCI by allowing the Virtual Environment (VE) to adapt to the user's emotional state.

Research on the difference in sensory perception between the real world and VR can be broadly divided into two perspectives: research from the perspective of illusions, and the other is from a purely cognitive perspective, including the cognitive load of the Working Memory (WM). Studies from the perspective of illusions have existed for a long time, including many on real-world phenomena. The most famous examples include the illusion phenomenon Rubber Hand Illusion (RHI) reported by Matthew, Jonathan [3], and others, the possession illusion and Proteus effect proposed by Yee and Bailenson [4].

As an extension of the RHI, Slater et al. [5][6] indicated that it could be produeced for virtual hands on a screen. Sanchez-Vives et al. [7] indicated that a visuo-haptic synchronization stimulus can induce a possession illusion for a virtual hand and suggested that this illusion can be induced without using a tactile stimulus.

Next, among studies about Proteus effects via avatar include, Yee and Bailenson showed that the use of avatars with different levels of attractiveness and height in appearance changes the way people communicate with others. Similarly, Oyanagi et al. [8] found that the use of a dragon avatar in a VR space can reduce the fear of heights. In a study by Tacikowski et al. [9], participants were exposed to images of the opposite sex's body through an HMD. Participants indicated that subjective

 ISBN: 978-1-68558-154-1

and implicit aspects of their gender identity, and stereotypical images of the opposite sex changed when they felt a sense of possession of the opposite sex's body.

Next, among studies on the relationship between VR and WM, Chiossi et al. [10] considered the influence of WM load, which leads to over- and under-stimulation, in the design of VR space. The authors designed an adaptive system to support the WM task execution based on electroencephalography (EEG) correlations between external and internal attention.

In parallel, a concept called by "Adaptive VR" has been discussed in recent years. Baker and Fairclough [11] described it as follows: Adaptive VR monitors human behavior, psychophysiology, and neurophysiology to create a real-time model of the user. This quantification is used to infer the emotional state of individual users and induce adaptive changes within the VE during runtime. Therefore, the authors argued that the efficacy of the emotional experience can be increased by modeling individual differences in the way users interact within a particular VE as a system.

To realize a seriese of practical studies, we need to follow the VR technological trends. Currently, many common VR technologies are in practical use. The low cost of HMDs, such as the Meta Quest2, has made it possible for consumers to easily experience VR content. There are two types of VR content: those in which the user does not move much and does move substantially in the VR space. Examples of the former include watching video content and browsing the web. In this case, the user's movements are mainly button operations and cursor movements using a controller, and the user rarely moves in both the real and VR spaces.

Examples of the latter include VR games and VR Social Networking Service (SNS) such as VR Chat [12]. VR games include those in which the user's actual body movements, such as swinging a sword or boxing, are synchronized with the movements of the avatar in the VR space, and those in which the user can move around in the VR space by operating a controller. Avatars are usually used in contents that allow users to move around in the VR space. Avatars are 3D objects that serve as the user's body in the VR space. Indeed, the use of avatars improves the realism of the VR experience and decreases the cognitive load in the VR space.

Given this background, the implementation of adaptive VR is becoming feasible. However, the main effects of its implementation on users should be understood based on a cognitive model. Studies have mainly focused on bottom-up content design with an awareness of adaptive VR. However, it is difficult for empirical developments to provide effects that create new phenomena. Hence, not only a bottom-up but also a top-down approach is necessary. As a stepping stone to this goal, we do the following in this study. We describe a model of the flow of information obtained by actual human perception through avatars in VR and the resulting human reactions, and confirm the validity of the user models proposed so far. The degree of immersion predicted because of the integration of multisensory information is also discussed. Understanding the role of multisensory information can enable us to design VR contents for individual users and how we can control sensory perception.

The remainder of this article is organized as follows. We describe the sense-perception cognitive model on VR in Section II. Next, we present an example of the difference between real-world and VR. Based on these, we finally explore the perception in the real and virtual worlds.

II. Description of the Cognitive Model for Sensory in VR

In general, physical information in the VR space is represented as follows. Objects in the VR space (VR objects) are represented by computer graphics, and their behavior is based on a program previously written to interact with the environment and other objects. The sound in the VR space is provided by artificially preparing audio data that is predicted in advance to be uttered in the space, and is played continuously in a background music-like manner, or by using a sound engine controlled by the user. Specifically, in the latter case, it can be attached to a VR object and played when certain conditions are met. Comprised of these elements, all human activities and virtual experiences in the VR space are performed by using the avatar as one's own body The avatar's movement is performed by tracking the user's real-world body movements. Tracking methods include three-point tracking, which consists of an HMD and two hand controllers, and full-body tracking, which uses motion capture and a tracking suit.

Consequently, the human experience in the VR space differs slightly from perception and cognition in the real world, and can be said to be the result of the interaction between avatar and VR objects, as well as the perception of the accompanying environment such as sound linked to these objects. Considering this, the model of human perception, cognition, and behavior in the VR space should be described with an awareness of the various interactions in the VR space with those in the real world.

- Perception of information
- Cognition of information
  - Attention
  - Memory
  - Decision
- Body motion based on perception
  - Human body motion
  - Interaction with VR objects

A. *Perception of information*

In general VR experiences using current HMDs, visual, auditory, and somatosensory information are used as perceptual information. The VR experience begins when the user puts on the HMD and views the images displayed on the lenses; by moving their head while wearing the HMD, the user can perceive the virtual space in the same way as they perceive the real world. Auditory information is output from the HMD's built-in or external speakers, and audio is played in response to the behavior of VR objects. The somatosensory information is

 ISBN: 978-1-68558-154-1

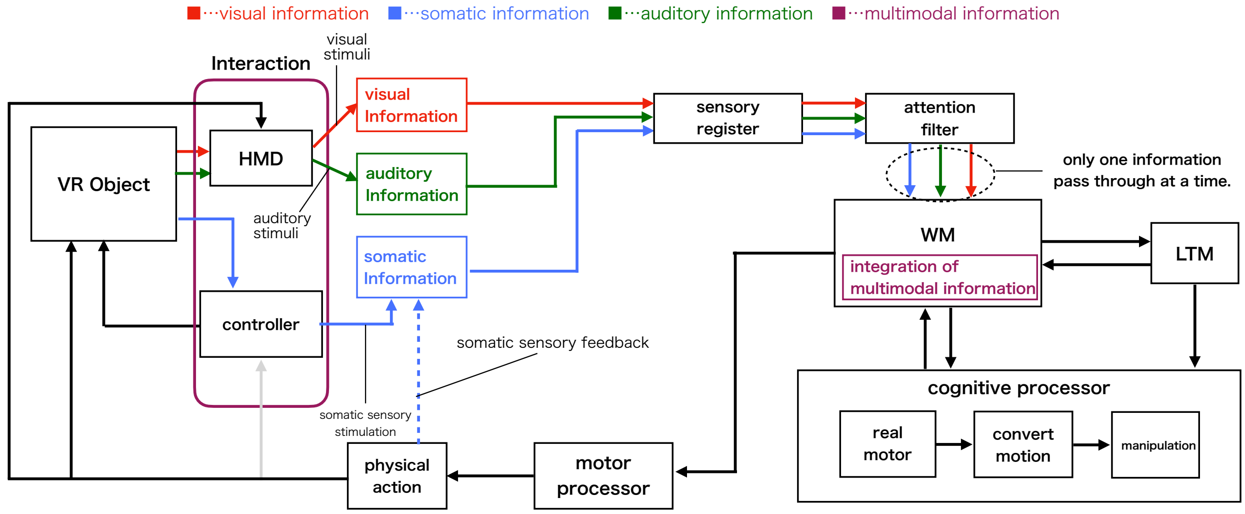

Figure 1. A Cognitive Model of VR Experience with HMD.

used to make operations in the VR space clearer by vibrating the controllers in both hands to generate tactile feedback when operating the User Interface (UI) in the VR space or selecting VR objects.

B. *Cognition of information*

1) Attention: Perceptual information moves to the sensory register, and then only the information to which the user's attention is directed passes through the selective filter and into the WM. Here, each sensory information does not completely enter the WM at the same time, but one piece of information passes through per processing.

2) Memory: If the sensory information obtained in the VR space is similar to that obtained in the real world, the user perceives the VR space as if it were a real space. In addition, based on the information in the Long-Term Memory (LTM), the user anticipates and expects the response of objects in the VR space to his or her actions, and engagement is generated.

3) Decision: Based on the perceptual information, the next action is determined. Here, when actions on a VR object are performed via a controller, the actions in the real world are converted into the corresponding controller operations.

C. *Body movement based on perception*

1) Human body movement: The operator (actual body) moves, and the avatar in the VR space moves in response to the movement. There are two methods for incorporating human motion into VR: (1) Image sensing by the camera attached to the HMD's basic UI operations (clicking and screen scrolling) and grasping VR objects (realized by holding something with a hand gesture) is possible. The high degree of synchronization between the actual hand and the avatar's hand motion is an advantage of this method. Conversely, precise manipulation, movements large enough to cause both hands to move out of the camera's field of view, and very fast hand movements are weaknesses. (2) Yaw, pitch, roll + relative position by controller. The accurate tracking of position, posture, and motion information by sensors is possible, and the sense of actual body motion is directly reflected during the operation, resulting in a high sense of immersion. However, if the reflection of body motion by the HMD is not synchronized with the actual body motion, it may cause a sense of discomfort and reduce the immersiveness of the VR experience.

2) Interaction with VR objects: VR objects not only appear to be three-dimensional, but can also be actually manipulated. Examples include playing a musical instrument or a push-button switch. Here, the immersiveness of the VR experience can be enhanced by providing not only a visual 3D effect, but also contextual information that one's actions affect the VR object.

D. *Integration of information*

Figure 2 shows the timeline of perceptual information in the WM when the perceptual information moves from the attention selector to the WM and activates information in the LTM from within the WM in Figure 1. Here, Information N refers to the information obtained from sensory organ N (e.g., vision). This information arrives in the WM at time t_N and exists for τ_N seconds. N is the number of perceived information. The long-term memory activated by these sensory information in the WM is denoted as h_M. In the case of Figure 2, there are two long-term memories activated by each of $N = 1, 2, 3$, and M is the number of i, j, k, l, p, and q in the WM. The time at which h_M arrives at the WM from the LTM is t_M, and the residence time is τ_M. Here, consider the contrast with perceptual information processing in the real space. For example, suppose that we now experience an event E_{rw} in real space. Suppose that H_i, h_j, h_k, h_l, and h_p of the LTM information are activated and processed simultaneously in the WM. Suppose that when an event E_{vw} close to E_{rw} is experienced in VR space, the information h_i, h_j, h_k, h_l, h_p in long-term memory is activated as a result of obtaining Information 1, 2, and 3, and processed simultaneously in the WM. Since the information processed in the WM is similar

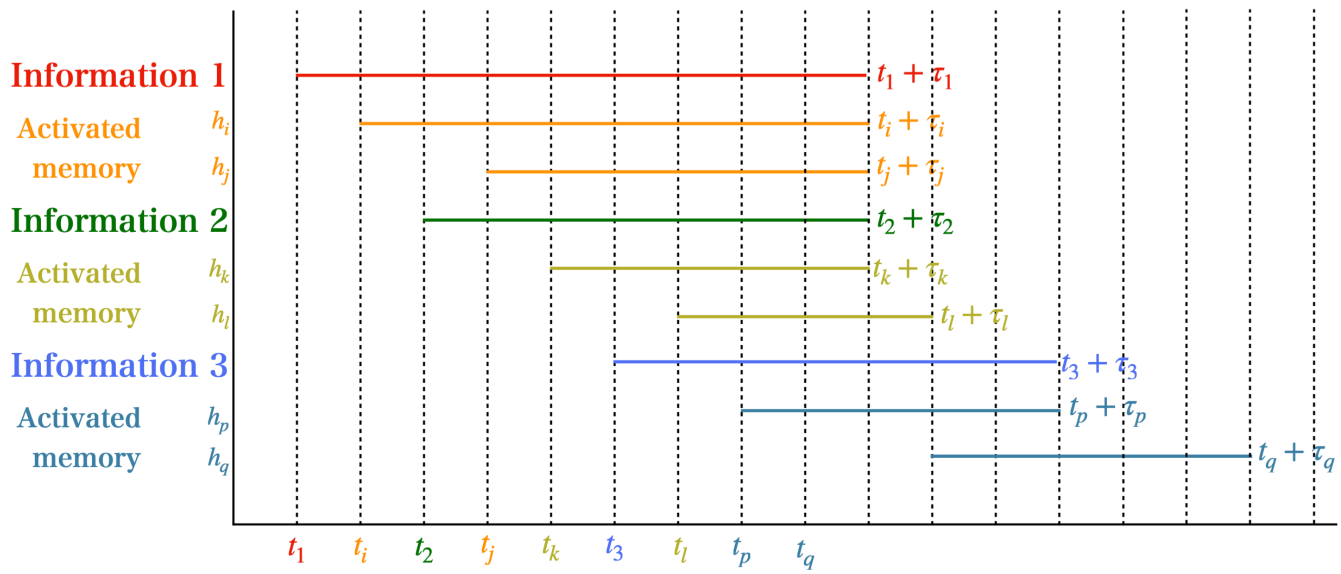

Figure 2. Staying timeline of sensory information stored in working memory and information invoked from long-term memory.

to that processed by the event E_{rw} in the real space, the VR experience is perceived as real and a sense of immersion is generated.

III. Examples of Discrepancies Between the Real and VR Worlds

A. Example 1: Playing Japanese Taiko

As an example, consider a situation in which the user plays *taiko* drums in a VR space. When the user sees a virtual *taiko* drum at time t_1, only visual information about the drum, I_1, exists in the user's WM. At time t_i, the sighting of the virtual *taiko* activates the *taiko* information h_i in the LTM, which becomes part of the information in the WM. Based on the user's experience in the real world, the user picks up the virtual stick and strikes the *taiko*, producing sound from the virtual *taiko*. At time t_2, the audio information reaches the user. Furthermore, the controllers of both hands vibrate, and somatosensory information reaches the user at time t_3. However, if the user knows from past experience that they feel air vibrations in their whole body when they hit the *taiko* drum, the information in the LTM does not match the information in the WM. This may cause a sense of discomfort and reduce the immersiveness of the VR experience.

B. Example 2: Communication within a VR space

Consider communication using avatars in a VR space. First, visual information, such as facial expressions and gestures of another avatar, exists in the user's WM at time t_1. Then, the voice of the other avatar reaches the user, and the voice information exists in the user's WM at time t_2. At this time, if the timing of the visual and audio information in the WM is off, such as if the other person's voice is heard from in front of the user even though the avatar of the conversation partner is behind the user, or if either of the two sensory information is unclear, the communication may feel uncomfortable or the immersiveness of the VR experience may be reduced.

IV. Perception in the Real/Virtual World

Based on Figures 1 and 2, we consider the perception of a phenomenon in the real (R) or virtual (V) space as follows. The chunk C_j stored in the LTM is constructed from the information group $I_i^{env}(t)$ obtained from sensory organ $i(1 \leq i \leq 5)$ in the past. Here, i refers to the five sensory organs possessed by a person. Each $I_i^{env}(t)$ passes through the attention filter $F_i^{env}(t)$ via the sensory register. And at time t, only the information obtained from a specific sensory organ passes through. C_j contains the information obtained from each sensory organ as a set $I(t)$ and is denoted as $C_j(I(t))$. Here, $I(t)$ is represented as follows:

$$I(t) = \{\boldsymbol{I} \mid I_i^{env}(t) F_i^{env}(t),\ 1 \leq i \leq 5\}.$$

The information that has passed through the attention filter is stored in the WM for a specific time, and a set of information $I(t)$ is sent to the LTM at the same time or with a time lag. In the LTM, $C_j(I(t))$ is matched with $C_j(I(t))$ based on the information in $I(t)$, and the closest or matching $C_j(I(t))$ is used as knowledge. The used knowledge is overwritten in the LTM through the WM in the form that the information in $I(t)$ is enhanced. Here, we target three sensory organs – visual, auditory, and somatic. We consider how the information flows through these three types of sensory organs in turn.

Suppose that at a certain time, a specific amount of information $I_i^{env}(t)(i = 1,\ 2,\ 3)$ is received from the external environment. $I_i^{env}(t)$ correspond to Information N in Figure

ISBN: 978-1-68558-154-1

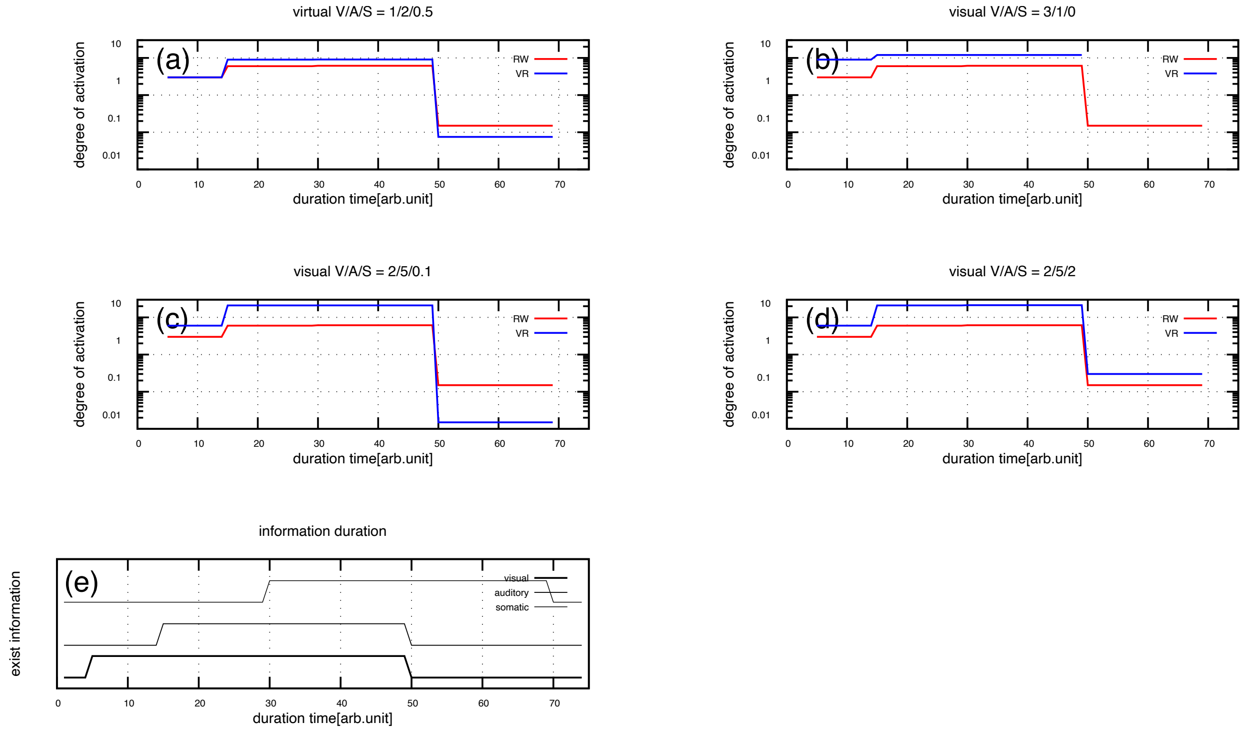

Figure 3. The trends of estimated $I^{syn}(t)$ which are changed three perceptual information(visual, auditory, somatic) amplified in Virtual Reality space.

2. Information N simultaneously activates several chunks. Although the degree of chunk activation varies, $I_i^{env}(t)F_i^{env}(t)$ is integrated into a single piece of information and sent to the LTM. In this case, the integration operator G can be used in various ways. The integrated information $I^{syn}(t)$ can be expressed as follows.

$$I^{syn}(t) = G(\ i,\ j,\ I_i^{env}(t)F_i^{env}(t),\ C_j(I(t))\ )$$

For the sake of simplicity, we simply add the amount of information and the degree of chunk activation as follows.

$$G^{env}(t) = \sum_i^n \sum_j^m I_i^{env}(t)F_i^{env}(t)C_j(I(t)) \qquad (1)$$

Figure 3 shows the trend of $I^{syn}(t)$ when the degrees to which visual, auditory, and somatic information are emphasized in VR are varied. The solid red line in the figure shows $I^{syn}(t)$ when visual, auditory, and somatic information are received in the real world. Here, we set $j = 1, 2$. Both visual and auditory information equal 1 for one, and 2 for the other. The somatic information is set to 0.5 on one side and 0.3 on the other. The solid blue lines indicate the degree to which the same information is distorted in VR.

Figure 3(e) shows the duration of information obtained from each sense. In contrast, Figures 3(a)$\sim$(d) shows the degree of integrated information activation calculated by Equation(1). Figure 3(a) shows the case where auditory is multiplied by a factor of 2 and somatic by a factor of 0.5. For $t < 50$, the VR space is slightly more chunk activated, but the characteristics are almost same. However, at $t \geq 50$, when only somatic information is perceived, the chunk activation in the VR space is lower. In Figure 3(b), the visual information is markedly increased, while the somatic information is not reproduced in the VR space. For $t < 50$, the activation of chunk in the VR space is markedly increased, but at $t \geq 50$, the somatic information is lost; Hence, there is no chunk activation in the VR space. In Figure 3(c), the somatic information is lowered to 0.1 and the information is emphasized in the form of visual<auditory. In particular, at $t \geq 50$, the somatic information is still present, but its effect is much smaller. Figure 3(d) is the case where the somatic information is also doubled. Compared with Figure 3(b) and (c), chunk activation remains high at $t \geq 50$.

The intensity of human sensation is expressed as a logarithm according to Weber-Fechner's Law. Therefore, as shown in Figure 3, even if the difference in sensory information is very slight, it suggests that the human senses can distinguish this difference. The sense of "slightly different from the real world" felt in VR content is thought to be caused by such slight differences in sensory information. The sensory information obtained in real space is not necessarily large, as shown in the example in Section III. However, it is easy to understand that these small differences lead to a sense of discomfort, which in turn indicates a decrease in immersive perception.

In the present case, we only dealt with a very simple integration of information. To advance our understanding of human sensory perception and use knowledge in VR spaces, scholars should develop a new approach that uses operators, such as Adaptive Control of Thought—Rational (ACT-R) [13] and Model Human Processor with Realtime Constraints (MHP/RT) [13] which incorporate Two Minds, to integrate information in a cognitive architecture [14][15][16].

V. Conclusion and Future Work

To realize adaptive VR, we need to design deeper immersion resulting from human interaction with real/VR spaces. As a first step, this study describes a sensory-cognitive model for VR spaces. The model is based on the integration of multimodal information, and the relationship between the three types of sensory information (visual, auditory, and somatic) and chunk activation. To understand the actual phenomena based on the described model, we consider and analyze the example of communication in the VR space with *taiko*, and refer to what kind of discomfort is likely to occur and its underlying causes. Connecting the two issues, multimodal information and chunk activatin, we undertake qualitatively research and explain the phenomenon that can occur when one or more types of information (visual, auditory, or somatic) is overemphasized or surpressed in a VR space. Expressing human sensory intensity as a logarithm according to Weber-Fechner's Law, we suggest that human senses can distinguish differences in sensory information, even if the differences are very slight. Considering these points, we are able to deepen our understanding of how the VR space realizes the immersive effect with impressive each other. Moreover, we are able to design "adaptive" immersive contents. In the future, it is necessary to investigate in experiments whether the degree of immersion felt by users changes when they experience VR content by changing the degree of emphasis of each sensory information. The metrics used to judge the degree of similarity between the real and virtual worlds can be defined as the overlap between the information held in the WM and the information in the LTM that has been activated up to that point in time. As the activation of information in the LTM is considered to be reflected in biological information, future experiments could be conducted using eye gaze and skin resistance measurements and subjective evaluation by means of questionnaires. Hysteresis can be considered based on the impact of inputs from the environment on the memory of the time series.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 20H04290, 22K12284, 23K11334, and National University Management Reform Promotion Project. The authors would like to thank Editage (www.editage.com) for the English language editing.

References

[1] A. Steed, Y. Pan, F. Zisch, and W. Steptoe, "The impact of a self-avatar on cognitive load in immersive virtual reality," in 2016 IEEE Virtual Reality (VR), 2016, pp. 67–76.

[2] S. M. H. Mousavi et al., "Emotion recognition in adaptive virtual reality settings: Challenges and opportunities," CEUR Workshop Proceedings, vol. 3517, jan 2023, pp. 1–20. [Online]. Available: https://sites.google.com/view/wamwb/

[3] B. Matthew and C. Jonathan, "Rubber hands 'feel' touch that eyes see," Nature, vol. 391, no. 6669, 02 1998, pp. 756–756. [Online]. Available: https://cir.nii.ac.jp/crid/1363388843275776640

[4] N. Yee and J. Bailenson, "The proteus effect: The effect of transformed self‐representation on behavior," Human Communication Research, vol. 33, 07 2007, pp. 271–290.

[5] M. Slater, D. Pérez Marcos, H. Ehrsson, and M. Sanchez-Vives, "Towards a digital body: the virtual arm illusion," Frontiers in Human Neuroscience, vol. 2, 2008. [Online]. Available: https://www.frontiersin.org/articles/10.3389/neuro.09.006.2008

[6] M. Slater, D. Pérez Marcos, H. Ehrsson, and M. Sanchez-Vives, "Inducing illusory ownership of a virtual body," Frontiers in Neuroscience, vol. 3, 2009. [Online]. Available: https://www.frontiersin.org/articles/10.3389/neuro.01.029.2009

[7] M. V. Sanchez-Vives, B. Spanlang, A. Frisoli, M. Bergamasco, and M. Slater, "Virtual hand illusion induced by visuomotor correlations," PLOS ONE, vol. 5, no. 4, 04 2010, pp. 1–6. [Online]. Available: https://doi.org/10.1371/journal.pone.0010381

[8] A. Oyanagi, T. Narumi, and R. Ohmura, "An avatar that is used daily in the social vr contents enhances the sense of embodiment," Transactions of the Virtual Reality Society of Japan, vol. 25, no. 1, 2020, pp. 50–59.

[9] P. Tacikowski, J. Fust, and H. H. Ehrsson, "Fluidity of gender identity induced by illusory body-sex change," Scientific Reports, vol. 10, no. 1, 2020, p. 14385. [Online]. Available: https://doi.org/10.1038/s41598-020-71467-z

[10] F. Chiossi, C. Ou, C. Gerhardt, F. Putze, and S. Mayer, "Designing and evaluating an adaptive virtual reality system using eeg frequencies to balance internal and external attention states," 2023.

[11] C. Baker and S. H. Fairclough, "Chapter 9 - adaptive virtual reality," in Current Research in Neuroadaptive Technology, S. H. Fairclough and T. O. Zander, Eds. Academic Press, 2022, pp. 159–176. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128214138000142

[12] VRChat Inc., "Vrchat," [Online]. Available from: https://hello.vrchat.com/, accessed, 2024.3.14.

[13] F. E. Ritter, F. Tehranchi, and J. D. Oury, "ACT-R: A cognitive architecture for modeling cognition," WIREs Cognitive Science, vol. 10, no. 3, 2019, p. e1488. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcs.1488

[14] M. Kitajima, Memory and Action Selection in Human-Machine Interaction. Wiley-ISTE, 2016.

[15] 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, 2012, pp. 41–58.

[16] 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, 2013, pp. 82–93.