Model of the Internet of Digital Education and its links to VR

Attila Kovari

Abstract—**Modern technologies are now essential to make the educational process more mobile, informative and versatile. The digitization of the educational institution, taking advantage of the opportunities provided by the Internet, provides an opportunity to exchange the accumulated experience and knowledge, and provides access to information for everyone. The mandatory introduction of online education due to COVID-19 has accelerated the development of digital education. The main direction in the development of education is moving towards modern online courses, which are already being used successfully. A comprehensive concept is needed to ensure that all IT used for learning, education and training is seamless, free of barriers and closed ecosystems. In connection with the above findings this article first introduces the model and of Internet of Digital Education (IoDE), describes the principles of IoDE, and then summarize the cognitive aspects and VR relations. The purpose of this article is to help readers recognize the IoDE as a whole and predict future trends in IoDE development.**

Index Terms—**Internet of Digital Reality; Digital Transformation; Education; Digital Pedagogy**

I. Introduction

As modern infocommunication technologies become more prevalent, we are moving towards a new digital world characterized by multi-modal entangled combination of IT with human and social cognitive systems [1]. The Internet and several 3D technology as virtual reality (VR), augmented reality (AR), mixed reality together with artificial intelligence (AI), and 5G networks are becoming ubiquitous technologies that affect society as a whole [2]. The impact of modern infocommunication technologies on the digital society is expected to lead to a qualitatively new kind of digital activities [3]. The concept of Internet of Digital Reality (IoD) as a set of network and related technologies for the management, of Digital Realities also have an impact on future forms of life, work and education [1].

As a result of the technologies mentioned above, education is undergoing major changes in recent decades and still is. Students now use different mental models to collect, process, and analyze data [4]. Greater emphasis will be placed on image-based information, scanned reading and responsive online communication [5][6][7]. They also use differen learning styles that include the ability to obtain information from groups and communities within online networking connections.

Attila Kovari is with the Institute of Digital Technology, Faculty of Informatics, Eszterházy Károly Catholic University, Eger, Hungary (e-mail: kovari.attila@uni-eszterhazy.hu) and GAMF Faculty of Engineering and Computer Sciences, John von Neumann University, Kecskemet, Hungary.

Modern technologies are now essential for the educational process to become more mobile, informative and versatile [8]. The digitization of the educational institution, using the opportunities provided by the Internet, provides an opportunity to exchange the accumulated experience and knowledge, and provides access to personalized information for all [9]. The digital transformation of universities means not only the introduction of IT solutions, but also the optimization of existing business processes, which makes them more flexible and "user-friendly" and can give the university greater popularity among students [10].

Presumably, in the near future, artificial intelligence will gain more and more ground in education, paper textbooks will largely disappear and the methodology of education will be transformed [11]. The main direction in the development of education is moving towards online courses, which are already being used successfully [12] [13]. A third major direction of technological change is simulators and virtual reality, which are becoming more widespread and can be used to teach a variety of skills and disciplines [14-20].

A comprehensive concept is needed to ensure that all IT used for learning [21-24], education and training is seamless [25-27], free of barriers and closed ecosystems. In connection with the above findings this article first introduces the importance and describes the principles of IoDE, present the model, then, and then summarize the cognitive aspects and VR relations.

For all IT tools used in learning, education and training to work seamlessly, without barriers and closed systems, a comprehensive approach is needed. The aim of the paper is to help readers to fully understand IoDE and to define the requirements of IoDE.

II. Importance and Principles of IoDE

The IoDE concept is a system for facilitating learning and teaching processes based on the interconnection of the internet and digital education. The term is based on the analogy of the Internet of Things (IoT) [28-33], which refers to the interconnection of various devices and systems via the Internet.

The importance of the concept can be interpreted in several ways:

Access: IoDE enables people around the world to access a variety of digital educational content, facilitating learning and skills development.

DOI: 10.36244/ICJ.2024.5.9

Flexibility: Teaching over the Internet makes the learning process more flexible and personalised, allowing learners to move at their own pace and choose topics that suit their interests.

Resource sharing: the IoDE facilitates the sharing of learning materials, resources and tools between teachers and students, thus increasing the quality and efficiency of education.

Interactive learning: Internet-based digital education gives students the opportunity to learn new skills in an interactive way, for example through virtual labs, games or simulations.

Connected world: the IoDE facilitates global collaboration in education, allowing students and teachers to share experiences and knowledge with each other, no matter where they are in the world.

In summary, the importance of the Internet of Digital Education concept lies in its ability to effective education, increase learning efficiency and enable global collaboration to disseminate knowledge. In the field of IoDE, the principles include efficiency, personalization, and a focus on self-education, which reflect the innovation trends of today's digital education that can be supported by the IoDE concept.

A. Effectiveness

It aims to introduce and apply new technological solutions to improve learning. The IoDE concept requires effective packaging and distribution of content between different actors. Effective IoDE based on individual approach to the overall organization of training, ensures interactivity, mobility and flexibility.

B. Put in context

Technology provides effective learning when we put things in context and it is led by a teacher. The use of technology needs to be placed in a context that meets the goals of education, we need to know how to find the answers using technology.

C. Consistency

An effective IoDE must offer a training ecosystem that provides a system of educational resources that combines the needs of students, universities, and companies interested in training so that the relationship between training opportunities is clear and traceable.

D. Internet and AI

The data processing capability of artificial intelligence and the connections provided by the Internet can bring many benefits to learning. For example, if as much data as possible is available, we can learn more from analyzing the information obtained from the analysis of that data. AI algorithms can extract and compare data from different learning environments, for example, to show which activities provide the most effective learning, the best results. These may include learning management systems, interactive learning environments, or even educational games, but not necessarily educational games that can enable data-rich learning through play. The importance of the information provided by the data is based on the fact that it can be used to plan and implement a personalized learning and teaching process that can optimize the level of competence of students with their learning activities.

E. Personalization

IoDE systems must ensure that the principle of one-size-fits-all content prevails over the principle of equal content for all learners, which is typical of the traditional education system. Adaptive learning systems driven by artificial intelligence provide an opportunity to consider formal and informal learning opportunities in addition to students' level of knowledge. The intelligent learning system adapts the teaching process and teaching materials to the individual learning speed of the students, highlighting the key concepts of the given curriculum and encouraging the learner to focus on the content of the curriculum that has not yet been mastered.

F. High level of self and AI assisted learning

The application of artificial intelligence in online learning has recently become an important new issue and method. Learning driven by artificial intelligence is of paramount importance in supporting self-directed learning, and this is partly consistent with teaching a human through machine control.

G. Interoperability

There is a need for standards that include requirements for providers of educational information technology programs that can ensure that technology is independent of devices and ecosystems and is fully interoperable [38]. This avoids that learners can only complete the learning activity in a given IoDE system, and effective learning is easily interoperable between systems, which is a prerequisite for the effective implementation of the IoDE concept. Some of the more significant standards related to learning are summarized in Figure 1.

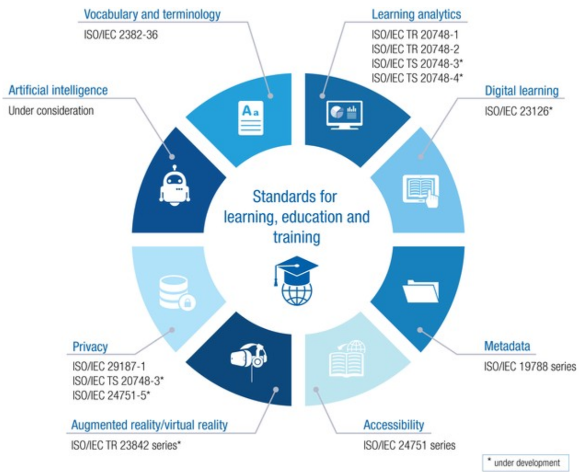


Fig. 1. Standards for learning, education and training [38].

H. Data security

In addition, data security and privacy must be guarantee for digital learning.

III. Generalized Model of IoDE

IoDE aims to provide a comprehensive approach to making teaching and learning processes more efficient and effective in the digital age, using the opportunities offered by the internet [34-36]. From the perspective of the IoDE model, an environment can be created for learners and educators to achieve personalized teaching and learning, taking into account individual goals, by interweaving the Internet and digital education. Based on the above, the IoDE model is shown in Fig. 2.

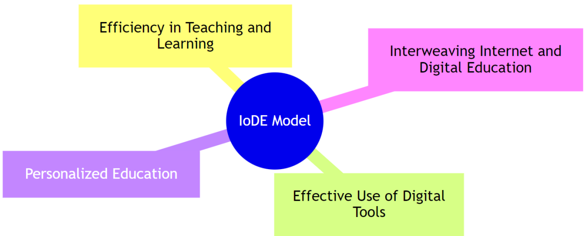


Fig. 2. Generalized model of IoDE

The model must reflect the methodologies shown in Figure 2, with particular regard to generations and learning styles. or early learners, IoDE can incorporate interactive and visually engaging content that aligns with their developmental stage. This includes intuitive interfaces, gamification, and multimedia resources that cater to their limited attention span and developing cognitive abilities. In case of school-age children, IoDE can offer more structured learning paths with a mix of interactive activities and educational content. Customization to accommodate varying learning speeds and styles is crucial, as is the integration of collaborative tools for peer interaction. For teenagers, IoDE can focus on more advanced, subject-specific content. This age group benefits from a blend of autonomy in learning and guided pathways. Tools for self-assessment, forums for discussion, and resources for project-based learning are particularly effective. Adult education through IoDE should focus on flexibility and relevance. This includes modular courses, career-oriented content, and resources that accommodate various life commitments. Adult learning often emphasizes practical, skill-based education that can be directly applied in the workplace.

IoDE can offer comprehensive curriculum and resources for parents and students engaged in home schooling. It provides access to diverse educational materials, interactive tools, and community forums for collaboration and support. In vocational training, IoDE can be instrumental in providing practical, skills-based learning through simulations, instructional videos, and interactive modules, catering to specific industry needs. IoDE enhances online education by offering a broad range of courses and resources accessible from anywhere. This includes MOOCs (Massive Open Online Courses), webinars, and virtual classrooms. IoDE plays a crucial role in making education more accessible and inclusive. It breaks down geographical and physical barriers, provides resources for learners with disabilities, and offers content in multiple languages to cater to a global audience.

As summarized above, a proposed detailed model for the Internet of Digital Education (IoDE) is built around the main elements and factors listed in Figure 3 and details are summarized in Table I.

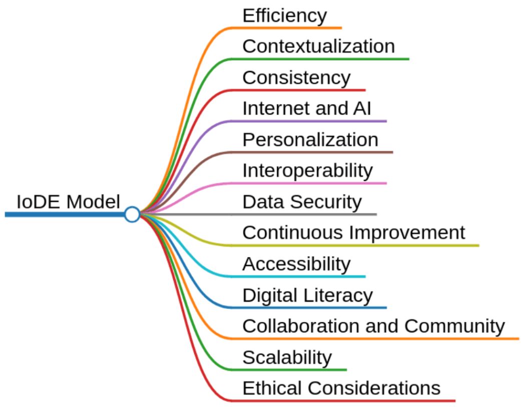


Fig. 3. Main factors of IoDE

TABLE I
Main elements and factors of IoDE

Factors	Details
Efficiency	Implementing and using new technologies to improve learning
	Use of new technologies to introduce new technologies and technologies for the delivery of training
Contextualization	Using technology in a context that is consistent with educational objectives
	Empowering teachers to help students find answers through technology
Consistency	Creating a training ecosystem that combines the needs of students, universities, and interested companies
	Clarity and traceability of links between training opportunities
Internet and AI	Leveraging the computing capabilities and benefits of AI and internet connectivity for learning
	Using AI algorithms to extract and compare data from different learning environments
	Using data to design and implement personalized learning and teaching processes
Personalization	Preference for personalized, universal content over content that is the same for all learners
	AI-driven adaptive learning systems that take into account formal and informal learning opportunities
	Adapting the learning process and teaching materials to individual learning speeds
	High levels of autonomous and AI-assisted learning
	AI-driven learning is paramount in supporting autonomous learning, partly in line with machine-driven instruction
Interoperability	Establishing prescriptive standards for education IT software providers

	Easy and efficient interoperability of effective learning between systems
Data Security	Ensuring data security and privacy for digital learning participants
Continuous Improvement	Regular updates and refinements of IoDE systems based on user feedback and technological developments
	Encouraging collaboration between educators, researchers, and technology providers
Accessibility	Ensuring IoDE systems are accessible to learners of different abilities, backgrounds, and learning preferences
	Application of inclusive design principles to promote equitable learning experiences
Digital Literacy	Teaching digital literacy skills to students, educators, and administrators
	Developing training programs and resources to support the development of digital literacy
	Incorporating AI-driven assessment and evaluation tools to measure student progress and learning outcomes
	Using data-driven insights to inform instructional design and curriculum development
Collaboration and Community	Building a sense of community between learners and educators in the IoDE environment
	Using technology to facilitate collaboration, communication, and peer support
Scalability	Designing IoDE systems to be scalable and adaptable to meet growing user numbers and changing needs
	Encouraging the use of IoDE solutions in different educational environments
Ethical Considerations	Addressing ethical issues related to AI in education, such as bias, fairness, and transparency
	Developing guidelines and rules to ensure responsible AI integration in the IoDE ecosystem

The IoDE model and concept brings several innovations to education compared to other concepts. IoDE develops education using digital technologies and the Internet, allowing learners and teachers to access educational materials and resources quickly and easily. Focusing education on collaboration. The IoDE concept can indirectly contribute, through the opportunities provided by the Internet, to enabling learners to progress adaptively at their own pace under flexible conditions. Taking advantage of global teaching-learning opportunities can increase efficiency [37-39].

IV. Cognitive Aspects of IoDE

The cognitive aspects of the Internet of Digital Education (IoDE) relate to the ways of learning, cognitive development and knowledge acquisition supported by technology and AI-driven solutions. The main cognitive aspects are:

A. Personalization

IoDE systems adapt to individual learning speeds, preferences and knowledge levels, ensuring a personalised learning experience.

B. Contextualization:

Technology is applied in a context that is consistent with educational goals, helping learners make meaningful connections and better understand complex concepts.

C. Interactivity

IoDE promotes active participation, allowing learners to interact with content and collaborate with peers and instructors, resulting in deeper understanding and improved problem-solving skills.

D. Independent learning

AI-assisted learning supports independent learning, helping students to achieve their own goals and define their own development path.

The importance of the cognitive aspects in the context of the Internet of Digital Education (IoDE) can be attributed to a number of factors. A personalized learning experience facilitates the individual development and progress of learners, taking into account their different abilities, interests and prior knowledge. A contextualization helps learners to understand and apply newly acquired knowledge to real-life situations, increasing the effectiveness of learning. Interactivity and active participation foster learners' critical thinking, problem-solving skills and deeper understanding of concepts. The independent learning promotes learners' self-activity and responsibility for their own learning process, which has a positive impact on motivation and achievement in the long term.

Taking cognitive aspects into account allows learning processes to be made more flexible to better adapt to learners' changing needs and circumstances. Equal access and equal opportunities: taking into account the cognitive aspects of IoDE promotes equal access and equal opportunities in learning by providing all learners with the opportunity to benefit from an optimal learning environment.

Overall, the importance of the cognitive aspects lies in their ability to enable more effective, personalised and flexible learning experiences, facilitating learners' success and progress in IoDE systems.

V. VR in the concept of IoDE

Virtual Reality can be used in the context of the cognitive aspects of IoDE as a tool to further enhance learning effectiveness, provide students with a deeper understanding and personalized learning experiences that support individual development. The linking of virtual reality (VR) to the concept of IoDE mainly focusing on the following aspects:

- Immersive learning experiences: VR allows learners to be fully immersed in a virtual environment that provides rich, interactive and realistic experiences during the learning process.
- Practical application: VR allows learners to try out and practise newly acquired knowledge in a safe and controlled environment, contributing to a more effective acquisition of skills and competences.
- Customizability and adaptability: VR environments are easily customizable and adapt to learners' individual needs, prior knowledge and learning pace, facilitating a personalized learning experience.

- Remote collaboration: VR enables learners and educators to collaborate and interact with each other despite any geographical distance, facilitating global collaboration and communication.
- Facilitated contextualisation: VR can be particularly useful in presenting complex or difficult-to-access concepts, phenomena and situations, helping learners to understand the context and apply knowledge to real-life situations.
- Maintaining motivation and interest: the exciting and interesting learning experiences provided by VR can increase learners' motivation and interest in the subject matter.

By combining VR and IoDE, it creates an immersive and interactive learning environment for both students and teachers, which significantly enhances the learning experience and the effectiveness of teaching. The introduction of VR technology on IoDE platforms allows for more visualisation of learning materials and increased learner engagement. The use of virtual reality in education facilitates distance learning and global collaboration, allowing students to access learning materials and tutor support anywhere and anytime. Combining VR and IoDE in education also contributes to supporting students with special needs. Instruction in virtual reality helps to better serve different learning styles and abilities, as well as offer learning experiences tailored to the individual needs of students. The combination of IoDE and VR allows for the continuous development and innovation of educational processes using new methods and tools. For educators, teaching in virtual reality opens up new opportunities for teaching and learning that more effectively encourage students to be creative, collaborative and problem-solving. Therefore, the link between virtual reality and the Internet of Digital Education can make a significant contribution to improving the learning experience, teaching methods and individual development in the field of education.

VI. Requirements for implementation of IoDE

The challenges of the 21st century require creative and innovative thinking and a changing division of labor between workers and machines as a result of the spread of artificial intelligence [39]. Digital competencies are gaining in value, bringing to the fore the knowledge, skills and attitudes that people can use to achieve their goals effectively [40][41]. Digital skills focus on the technological, communication, information and multimedia aspects that lead to consciously effective action to create value [42] [43].

Important changes are taking place in the areas of intelligent manufacturing, fully virtualized interfaces, and remote control devices. Future professionals need to have different competencies than before, so they need to be developed. The increasingly complex interaction of people with modern technologies will bring significant changes. The field of cognitive infocommunication [44] will become even more important, in which digital competencies problem-solving, creativity and collaboration [45-48], effective communication [49] and study results [50], emotional intelligence [51-53], BigData [54], 3D [55], VR [56] or effective use of modern digital devices [57-59]. In addition, human-computer interfaces (HCIs) such as eye-tracking [60-62] or brain-computer interfaces (BCIs) [63] can be an alternative possibility to analyze the effectiveness of different cognitive processes.

The need to combine advanced technological and soft skills requires the use of educational environments and methodologies that are based on the integration of multiple disciplines, allowing the emergence of IoDE applications. The application of IoDE can effectively improve digital development of competencies in the Internet, learning by integrating spaces and educational experiences that can lead to effective competence development.

The most basic digital competencies of current learners have already developed to a greater extent, as they have grown up in the digital environment since childhood. Virtual online learning environments are already available to students and many technological opportunities are not yet fully exploited, we are witnessing an incredible transformation of educational environments that lay the foundations for digital educational institutions and digital universities. The days, activities, communication and interactions of a significant part of the students and already the teachers with the outside world take place largely through the Internet. Learning styles are also changing and changing, with students now using different methods of data collection and processing than they did 10-20 years ago. Image and video-oriented information, scanned reading, and information from multiple sources that have developed a different learning style in students come to the fore. These all predict the effective use of IoDE.

The development of ICT enables the introduction of many innovative teaching techniques and methodologies in distance learning, also in the case of flipped classroom or hybrid learning techniques.

It can also be seen from the above that an essential condition for effective IoDE is that the technology is available at the right level, in the right number and with the right support in the educational institution. Of course, this is not available in all institutions and at the same level, not only in terms of digital technology tools, but also in terms of the knowledge, methodology, useful experience, ideas and elaborated curricula needed to use them. Understanding of digitization and its transformative effects, developing and integrating a digital switchover strategy, and exploiting the potential of IoDE are key factors in implementing effective IoDE.

In summary, the main requirements for effective IoDE implementation are:

- in general, the institution should be aware of digitization and its transformative effects, understand the role of digitization and provide appropriate support;

- the institution should plan ahead, its application should be an important goal in its vision and strategic plan, the staff and faculty of the institution should be involved and support the transformation, and proper communication is essential;

- the institution should support those at the forefront of transformation, help and support the testing of new and innovative working methods, including the provision of IT professional support;

- take advantage of the systemic and community-level opportunities provided by IoDE to increase the effectiveness of education and to adapt to modern forms of learning;

- students and staff should be kept up to date with the technological and methodological potential of IoDE.

VII. Conclusion

The nature of education known today is facing a significant change due to the emergence of advanced technologies. Higher education institutions compete for students, play an important role in this competition, and have the potential to benefit from advanced digital communication technologies based on the Internet and methodologies based on them in education.

University leaders need to recognize that there is an increasing need to modernize their education systems and processes, and to transform their pedagogical, teaching methodology, and curriculum, subject, and thematic practices. Institutions develop unique plans and strategies, try to introduce unique digital solutions in order to take advantage of the opportunities provided by advanced technologies, but the most effective solutions do not necessarily come to the fore during the implementation. Resources devoted to the development of ICT systems do not always bring the desired results. The present study approaches the exploitation of the Internet-based potential of digital education through the conceptual definition of the digital framework introduced with the introduction of IoDE.

IoDE ecosystem connects students, teachers, and other cognitive entities, digital education systems using the combinations of most relevant technologies and data that create a higher-level functional integration to achieve higher-level learning efficiency. This could be one of the foundations for the successful educational institutions of the future.

References

[1] P. Baranyi, Á. Csapó, T. Budai, Gy. Wersényi, "Introducing the Concept of Internet of Digital Reality – Part I," *Acta Polytechnica Hungarica*, vol. 18, no. 7, pp. 225–240, 2021, DOI: 10.12700/APH.18.7.2021.7.12.

[2] J. Orlosky, K. Kiyoshi, T. Haruo, "Virtual and augmented reality on the 5G highway," *Journal of Information Processing*, volt 25, pp. 133–141, 2017. DOI: 10.2197/ipsjjip.25.133

[3] T. Lengyel Molnár, "Library strategies in the Cyber-Physical Society," *Proceedings of Society for Information Technology & Teacher Education International Conference 2020*, Waynesville (NC), Association for the Advancement of Computing in Education, pp. 1048–1054, 2020.

[4] K. Chmielewska, "From Mathability to Learnability", *Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 513–516. DOI: 10.1109/CogInfoCom47531.2019.9089891

[5] K. Chmielewska, "Towards' Learnability' – What we benefit from mathability", *Acta Polytechnica Hungarica*, vol. 19, no. 1, pp. 171–192, 2022. DOI: 10.12700/APH.19.1.2022.19.11

[6] K. Chmielewska, A. Gilányi, "Educational context of mathability", *Acta Polytechnica Hungarica*, vol. 15, no. 5, pp. 223–237, 2018. DOI: 10.12700/APH.15.5.2018.5.13

[7] K. Chmielewska, A. Gilányi, A. Łukasiewicz, "Mathability and mathematical cognition", Proceedings of the 7th IEEE Conference on Cognitive Infocommunications, 2016, pp. 245–250. DOI: 10.1109/CogInfoCom.2016.7804556

[8] R. Racsko, E.M. Troll. "Our digital education habits in the light of their environmental impact: the role of green computing in education," *Teaching Mathematics And Computer Science*, vol 18 no. 1, pp. 69–86, 2020.

[9] I. Horváth, "An Analysis of Personalized Learning Opportunities in 3D VR," *Frontiers Computer Science*, vol 20, ArtNo: 673826, 2021. DOI: 10.3389/fcomp.2021.673826

[10] M. Rajcsányi-Molnár, S. Czifra, "Development of an integrated online education system at the University of Dunaújváros," *Fenntarthatósági terek*, DUE Press, 2021, pp. 84–101.

[11] P. Négyesi, I. Oláh Téglási, R. Racsko, "The application of a custom-made mathematical software in order to support problem-solving strategy of high school students," *Proceedings of XXXIII. DidMatTech 2020 Conference, New Methods and Technologies in Education in Research and Practice*, Budapest, Hungary, 2020, pp. 402–407.

[12] B. Orosz, C. Kovács, D. Karuovic, Gy. Molnár, L. Major, V. Vass, Z. Szűts, Zs. Námesztovszki, "Digital education in digital cooperative environments," *Journal of Applied Technical and Educational Sciences*, vol 9, no 4, pp. 55–69, 2019. DOI: 10.24368/jates.v9i4.149

[13] K. Chmielewska, W. Ciskowska, D. Glazik, D Marcinek, K. Wojciechowska, A. Gilányi, "Learnability – are we ready for distance learning?", *Proceedings of the 11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 465–470. DOI: 10.1109/CogInfoCom50765.2020.9237881

[14] M. Dergham, A. Gilányi. "Application of virtual reality in kinematics education", *10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 107–111. DOI: 10.1109/CogInfoCom47531.2019.9089971

[15] A. Gilányi, A. Rácz, A. M. Bólya, K. Chmielewska, "Early history of Hungarian ballet in virtual reality", *10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 193–198. DOI: 10.1109/CogInfoCom47531.2019.9089943

[16] J. Katona, "A Review of Human–Computer Interaction and Virtual Reality Research Fields in Cognitive InfoCommunications," *Applied Sciences*, vol. 11, no. 6, pp. 2646, 2021. DOI: 10.3390/app11062646

[17] A. Kovari: Possibilities of applying project education in IT education. In *XXVI. International Conference on Computer Science and Education*, 2016, pp. 247–251.

[18] M. Dergham, A. Gilányi, "On a system of virtual space for teaching kinematics", *10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 411–414. DOI: 10.1109/CogInfoCom47531.2019.9089935

[19] A. Gilányi, A. Rácz, A. M. Bólya, J. Décsei, K. Chmielewska, "A presentation room in the virtual building of the first National Theater of Hungary", *11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 517–522. DOI: 10.1109/CogInfoCom50765.2020.9237828

[20] A. Rácz, A. Gilányi, A. M. Bólya, J. Décsei, K. Chmielewska, "On a model of the first National Theater of Hungary in MaxWhere", *11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 573–574. DOI: 10.1109/CogInfoCom50765.2020.9237848

[21] B. Holmberg, *Theory and Practice of Distance Education, Routledge*, 2005.

[22] Gy. Molnár, "Teaching and Learning in modern digital Environment," *Proceedings of the IEEE 13th International Symposium on Applied Machine Intelligence and Informatics*, pp. 213–217. DOI: 10.1109/SAMI.2015.7061878

[23] E. Gőgh, R. Racsko, A. Kővári, "Experience of Self-Efficacy Learning among Vocational Secondary School Students," *Acta Polytechnica Hungarica*, vol. 18 no. 1, pp. 101–119, 2021. DOI: 10.12700/APH.18.1.2021.1.7

[24] T. Lengyelné Molnár, "Changing reading habits and methodological options resulting from digital transformation," *Journal of Applied Technical And Educational Sciences*, vol. 9, no. 4, pp. 27–42, 2019. DOI: 10.24368/jates.v9i4.108

[25] M. Rajcsányi-Molnár, A. Bacsa-Bán, "From the Initial Steps to the Concept of Online Education" – Teacher Experiences and Development Directions Based on Feedback from Online Education Introduced During the Pandemic. *Central European Journal of Educational Research*, vol. 3, no 3, pp. 33–48, 2021. DOI: 10.37441/cejer/2021/3/3/9612

[26] Gy. Molnár, Zs. Namestovski, R. Pinter, "Massive Open Courses (MOOCs) held in Serbian," *Journal for Information Technology Education Development and Teaching Methods of Technical and Natural Sciences*, vol 6, no 1, pp. 12–16, 2016.

[27] D. E. Salinas-Navarro, C. L. Garay-Rondero, "Experiential learning in Industrial Engineering education for Digital Transformation," *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, 2019, pp. 1–9, DOI: 10.1109/TALE48000.2019.9225984.

[28] K. Thorsten, R. Van Kranenburg, S. Lange, "Introduction to the Internet of Things," *Enabling Things to Talk*. Springer, Berlin, Heidelberg, 2013, pp. 1–10. DOI: 10.1007/978-3-642-40403-0_1

[29] L. Srivastava, T. Kelly, "The internet of things," International Telecommunication Union, Tech. Rep 7, 2005.

[30] T. Guzsvinecz, T. Medvegy, V. Szucs, "A brief review on the challenges of Internet of Things and their solutions," *Proceedings of the Pannonian Conference on Advances in Information Technology (PCIT'2019)*, 2019, pp. 181–186.

[31] E. Harald, P. Goodridge, J. Haskel, "The Internet of Things and economic growth in a panel of countries," *Economics of Innovation and New Technology*, vol 30, no 3, pp. 262–283, 2021. DOI: 10.1080/10438599.2019.1695941

[32] Cisco, "The Internet of Everything - Cisco IoE Value Index Study." 2013. https://www.cisco.com/c/dam/en_us/about/business-insights/docs/ioe-value-index-faq.pdf

[33] Gy. Wersényi, Á. Csapó, T. Budai, P. Baranyi, "Internet of Digital Reality: Infrastructural Background – Part II," *Acta Polytechnica Hungarica*, vol. 18, no. 7, pp. 225–240, 2021, DOI: 10.12700/APH.18.8.2021.8.5.

[34] J. Lengsfeld, *Digital Era Framework*, 2019. https://joernlengsfeld.com/files/digital-era-framework-en.pdf

[35] Z. Balogh, C. Klimeš, "Modelling of education process in LMS using Petri nets structure," *Proceedings of the IADIS International Conference e-Learning, MCCSIS*, Freiburg, 2010, pp. 289–191.

[36] P. Porteleki, "IoT integration in Microsoft Dynamics ERP," Computers & Learning, vol 3, no 1, pp. 1–11, 2020.

[37] S. Oza *et al*., "IoT: The Future for Quality of Services," In: ICCCE 2019. Lecture Notes in Electrical Engineering, vol 570. Springer, Singapore, 2020, DOI: 10.1007/978-981-13-8715-9_35
https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution#/media/File:Internet_of_Things_(IoT).png

[38] A. Price, "Why the future of learning is digital, Digital technologies are transforming learning and building a more sustainable future," e-tech, https://medium.com/e-tech/advancing-the-digitization-of-education-8e47362b2165

[39] World Economic Forum, "The future of jobs report 2018," Geneve: WEF, 2018. https://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

[40] Gy. Molnár, D. Sik, "The virtual toolkit of digital instruction and its application in digital work forms," *Proceedings of the 2020 11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 599–602. DOI: 10.1109/CogInfoCom50765.2020.9237855

[41] Gy. Molnár, Z. Szűts, Z. Balogh, "Modern digital Web 2.0 devices and services supporting the teaching of Technology and Informatics," *Proceedings of the IEEE 17th International Symposium on Intelligent Systems and Informatics Proceedings*, 2019, pp. 89–94. DOI: 10.1109/SISY47553.2019.9111555

[42] T. Guzsvinecz, J. Szűcs, "Using Analytics to Identify When Course Materials Are Accessed Relative to Online Exams during Digital Education", *Education Sciences*, vol. 11 no. 10, p. 576, 2021. DOI: 10.3390/educsci11100576

[43] Gy. Molnár, K. Nagy, Z. Balogh, "The role and impact of visualization during the processing of educational materials, presentation options in education and in the virtual space," *Proceedings of the 2019 10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 533–537. DOI: 10.1109/CogInfoCom47531.2019.9089917

[44] D. Mattyasovszky-Philipp, B. Molnár, "An Architectural Approach to Cognitive Information Systems," *Acta Polytechnica Hungarica*, vol 17, no 2, pp. 237–253, 2020. DOI: 10.12700/APH.17.2.2020.2.13

[45] G. V. Kövecses, B. Lampert, A. Pongrácz, and I. Lőrincz, "University Lecturers' Distance Learning Experiences Gained during the COVID-19 Pandemic Period," *Proceedings of the 2020 11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 325–330. DOI: 10.1109/CogInfoCom50765.2020.9237890

[46] T. Guzsvinecz, É. Orbán-Mihálykó, E. Perge, and C. Sik-Lányi, "Analyzing the Spatial Skills of University Students with a Virtual Reality Application using a Desktop Display and the Gear VR," *Acta Polytechnica Hungarica*, vol. 17, no. 2, pp. 35–56, 2020. DOI: 10.12700/APH.17.2.2020.2.3

[47] Z. R. Rajtik, "Vocational training 4.0 strategy ideas for the field of mechanical engineering," *Transactions on IT and Engineering Education*, vol 3, no 1, pp. 16–31, 2020.

[48] M. Rajcsányi-Molnár, "MaTech: Organizing a creative math competition based on digital tools for secondary school students," *Felsőoktatási innovációk a tanulás korában*, MELLearN, 2019, pp. 19–32.

[49] D. Mattyasovszky-Philipp, A. M. Putnoki, B. Molnár, "The Unrepeatable Human Mind – Challenges in the Development of Cognitive Information Systems – What Makes a Machine Human?," *Electronics* 2022, vol 11, ArtNo 394. DOI: 10.3390/electronics11030394

[50] Z. Balogh, M. Kuchárik, "Predicting Student Grades Based on Their Usage of LMS Moodle Using Petri Nets," *Applied Sciences,* vol 9, no 20, ArtNo: 4211, 2019. DOI: 10.3390/app9204211

[51] G. V. Kövecsesné, A. Benyák, A. Makkos, and Á. Fehér, "Emotional intelligence development in online education," *Proceedings of the 12th IEEE International Conference on Cognitive Infocommunications*, 2021, pp. 517–521.

[52] M. Magdin, *et al*., "Automatic detection and classification of emotional states in virtual reality and standard environments (LCD): comparing valence and arousal of induced emotions," *Virtual Reality,* vol 25, no 4, pp. 1029–1041, 2021. DOI: 10.1007/s10055-021-00506-5

[53] A. Makkos, Á. Fehér, A. Mészáros, G. V. Kövecsesné, A. Benyák, "Examining the emotional intelligence and coping strategies of trainee teachers and in-service teachers," *ICFE 2021, School and teacher education* 2030, 2021, p 18.

[54] A. Benedek, Gy. Molnár, Z. Szűts, "Practices of Crowdsourcing in relation to Big Data Analysis and Education Methods," *Proceedings of the IEEE 13th International Symposium on Intelligent Systems and Informatics*, 2015, pp. 167–172. DOI: 10.1109/SISY.2015.7325373

[55] Z. Szűts, "An Iconic Turn in Art History – The Quest for Realistic and 3D visual Representation on the World Wide Web," *The Iconic Turn in Education (Visual Learning)*, Frankfurt am Main, Peter Lang, 2012, pp. 59–66.

[56] I. Horváth, "Innovative engineering education in the cooperative VR environment," *Proceedings of the 7th IEEE International Conference on Cognitive Infocommunications*, 2016. pp. 359–364. DOI: 10.1109/CogInfoCom.2016.7804576

[57] G. V. Kövecsesné, "Supporting the learning of teaching students with digital tools," *Journal of Applied Technical and Educational Sciences*, vol. 10, no. 4, pp. 105–124, 2020. DOI: 10.24368/jates.v10i4.208

[58] V. Szücs and C. Steiner, "Multifunctional Chatbot for search support and secure the host," *Proceedings of the 12th IEEE International Conference on Cognitive Infocommunications*, 2021, pp. 943–948.

[59] I. Horváth, "MaxWhere 3D Capabilities Contributing to the Enhanced Efficiency of the Trello 2D Management Software", *Acta Polytechnica Hungarica*, Vol. 16, No. 6, pp. 55–71. 2019. DOI: 10.12700/APH.16.6.2019.6.5

[60] J. Katona, "Measuring Cognition Load Using Eye-Tracking Parameters Based on Algorithm Description Tools," *Sensors*, vol. 22, no. 3, pp. 912, Jan. 2022. **DOI**: 10.3390/s22030912

[61] J. Katona, "Analyse the Readability of LINQ Code using an Eye-Tracking-based Evaluation," *Acta Polytechnica Hungarica*, vol. 18, no. 1, pp. 193–215, 2021. **DOI**: 10.12700/APH.18.1.2021.1.12

[62] J. Katona, "Clean and Dirty Code Comprehension by Eye-tracking Based Evaluation using GP3 Eye Tracker," *Acta Polytechnica Hungarica*, vol. 18, no. 1, pp. 79–99, 2021. **DOI**: 10.12700/APH.18.1.2021.1.6

[63] J. Katona, "Examination and comparison of the EEG based attention test with CPT and TOVA," in *15th IEEE International Symposium on Computational Intelligence and Informatics*, Budapest, Hungary, 2014, pp. 117–120. **DOI**: 10.1109/CINTI.2014.7028659

Attila Kovari is Professor at Eszterhazy Karoly Catholic University, University of Dunaujvaros, Obuda University and John von Neumann University, Hungary. He holds a Ph.D. degree in Engineering and Habilitation in Computer Engineering and Education. His research areas are human-computer interaction, brain-computer-interfaces, eye-tracking analysis, IT in industrial applications.

Special Issue
of the Infocommunication Journal

Textual Analysis of Virtual Reality Game Reviews

Tibor Guzsvinecz, and Judit Szűcs

Abstract—**Virtual reality systems are complex and made of various parts. Since a person is an integral component of such systems, virtual reality technologies also have a cognitive aspect. As such, these technologies engage with the perceptual, attentional, and decision-making processes of users. Consequently, they can be considered cognitive tools. Thus, it is imperative to understand what people think of such environments. To take a step towards this understanding, textual reviews of virtual reality games made for entertainment were investigated using text mining methods. Thus, 1,635,919 textual reviews were scraped from the Steam digital video game distribution platform in the spring of 2023. The reviews were grouped by whether they were positive or negative. According to the results, the following conclusions can be made regarding virtual reality games: 1) Negative reviews are significantly longer than positive ones. 2) Negative reviews are written significantly earlier than positive ones, although no correlation was found between the review type and the playtime before writing the review. 3) The most frequent words and word correlations are different between review types since negative reviews are more focused on game mechanics and bugs. Due to the results, insights can be provided to virtual reality game developers to help them refine their games.**

Index Terms—**computer games; game review; player experience; Steam; textual analysis**

I. Introduction

In this digital age, the popularity of virtual reality (VR) applications is unquestionable. As VR is a synthetic reality, it is possible to create virtual environments (VEs) that are either similar to or different from the real world. Depending on the goal of such virtual spaces, they could be created for various purposes [1, 2]. For example, VR technologies play an important role in education [3–6], training [7, 8], healthcare [9, 10], and even entertainment [11, 12].

As can be expected based on these previously mentioned fields, a VR system requires a person to interact with it to work perfectly [13]. This is due to the fact that these systems have cognitive aspects as well [14, 15]. VR can also be considered a cognitive tool [16]. Therefore, when designing such systems, it is important to keep the target groups in mind during the process. Naturally, after a VR system or application is implemented, users can experience it and leave feedback about its strengths and weaknesses. This is also the case with VR applications for research or entertainment purposes. Feedback can also be verbal or textual.

Understanding textual feedback is crucial to know how people reacted to various experiences. This textual feedback can come in the form of a review [17, 18], which allows game developers to learn how to improve their games [19]. Video game reviews can contain several topics such as achievements, accessories, general experience, social interaction, social influence, narrative, visual/value, and information about bugs in the game [20]. Thus, it is possible to know how playable a game is from the reviews [21]. Reviews can also detail several factors that can make a game popular [22]. They can also serve as a product review: they can influence other players whether to buy or not to buy a game [23, 24]. Also, critics tend to highlight different aspects of games than players [25].

As can be seen, reviews contain critical information about the experiences of players. Furthermore, many reviews are written daily, thus it is possible to analyze a large number of data [26]. Due to this data, game developers and researchers can ascertain the strengths and weaknesses of the analyzed games. Naturally, these can improve the quality of future games [27, 28].

Thus, the goal of this paper is to take a step towards understanding VR games and applications made for entertainment using textual analysis methods. For this, the following research questions (RQs) were formed:

- RQ1: How long are the reviews of VR games?
- RQ2: Is there a correlation between playtime and the length of VR game reviews?
- RQ3: What are the most frequent words and their word associations in the reviews?

The first question aims to investigate the length of textual reviews for VR games, comparing word count of positive and negative reviews. The second one seeks to explore whether there is a relationship between playtime and the length of reviews. By examining potential correlations, it is possible to determine if users who spend more time playing VR games are more likely to write longer reviews, and vice versa. Lastly, by answering the third one we can uncover key terms, expressions, or themes that emerge frequently in both positive and negative reviews. Thus, understanding reviews can provide insights into the level of detail and depth of user feedback, which can be valuable for developers and researchers.

Submitted on...
Tibor Guzsvinecz is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: guzsvinecz.tibor@zek.uni-pannon.hu). He is the corresponding author.

Judit Szűcs is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: szucs.judit@zek.uni-pannon.hu)

DOI: 10.36244/ICJ.2024.5.10

Therefore, this paper is structured as follows. Section 2 details the Steam scraping process and the textual analysis. The results are presented in section 3. In section 4, the results are discussed along with the limitations of the study. Lastly, conclusions are made in section 5.

II. Materials and methods

The reviews on the Steam digital game distribution platform were chosen for the analysis. The platform was chosen because it was the largest digital game distribution platform in 2017 [29]. In 2021, it was still one of the biggest platforms with more than 50,000 video games in its library [30]. Not to mention, due to the Covid-19 pandemic, digital video game distribution platforms gained popularity as well [31].

On Steam, players have the opportunity to write reviews on a game's page, although they have to register a free profile and play the said game beforehand. Naturally, it is free to browse and read reviews on the platform. Valve Corporation – the creator of this platform – has developed an application programming interface (API) to freely access and scrape some of the platform's contents, including video game reviews [32].

Thus, this section presents the following. In the first subsection, the reviews on Steam are defined. Afterward, the scraping process is detailed in the second subsection. Lastly, the textual analysis is shown in the third subsection.

A. Video game reviews on Steam

On the Steam digital video game distribution platform, each game has its own store page. These pages contain the reviews as well. According to the Steam API, each review object has multiple components that can be accessed [33]. The following components were used for this research: the textual part, the language of the review, playtime when writing the review, and whether a review is positive or not. For the latter, it should be mentioned that Steam does not have a numerical rating system. Thus, players can either recommend the game or not. The first represents a positive review, while the latter is a negative one. Naturally, in the previously mentioned textual component, players can write down their experiences with no character limitation.

B. The scraping process

The statistical program package R was used for the scraping process along with its *rvest* and *httr* packages. The Steam API was also used to access content during the process. The scraping process was conducted in the spring of 2023.

First, the list of games and applications was created using the *GetAppList* function that could be found in the Steam API. It returned the list of all application IDs on the platform. Then, the *read_html* function was used on every application ID to load their store's pages which contain tags. Among other things, these tags allow us to see whether a game is a VR application. The IDs of the VR application were saved into a vector.

Then, the actual scraping started. According to the documentation of Steam, the *GET* function was used in R [34]. Due to the large amount of data, only English and the most recent 1,000 reviews were scraped per game. Naturally, if a game did not have at least 1,000 reviews, all were scraped. Overall, 1,635,919 reviews were scraped. Out of them, 79.02% were positive and 20.98% were negative.

C. Data analysis

After the scraping was finished, the data analysis process started. First, the number of words was calculated with the regular expression of '[\\w\']+'. This also required the *stringr* package, and the *str_count* function in it. The number of words was also grouped by review type.

To count the occurrences of each word, the *tm* package was used. First, a corpus was created of the textual parts found in the reviews. Then, the text was manipulated as follows: words were converted to lower cases; English stopwords, punctuation, and numbers were removed; whitespaces were stripped; and lastly, the words were stemmed using Porter's algorithm [35]. Afterward, the word occurrences could be calculated. The *ggplot2* package was used to plot the most frequent words.

When investigating the word correlations with frequent words, the *tidyverse*, *tidytext*, *widyr*, *igraph*, and *ggraph* packages were used. After empirical testing, the following parameters were used. First, the words had to occur at least once in at least 100 reviews per game. Then, by using pairwise correlation, the coefficients were assessed among all possible combinations. For plotting, in the case of positive reviews the correlation coefficient was set to at least 0.7, while in the case of negative reviews, it was set to at least 0.6. Otherwise, the figures would have become unreadable.

Letter-value plots were used to better visualize the distribution of data. These types of plots were designed for large datasets and they allow us to see more reliable estimates beyond the quartiles [36]. After them, the upper and lower eights are shown, then upper and lower sixteenths, and so on until the process reaches a stopping criterion. A 95% confidence interval is determined around each letter-value. If an overlap is found between this confidence interval and the previous letter-value, then the current letter-value and the following values are not plotted anymore. Each interval is shown with darker and smaller boxes. Due to the large number of data, a graphical method was used to check whether each dataset followed Gaussian distribution. Thus, quantile-quantile (Q-Q) plots were used. Due to the results found in the previously mentioned plots, the non-parametric Wilcoxon rank sum test was used when comparing word numbers and playtimes between review types [37]. Similarly, when assessing the correlation between playtime and review length, the Spearman rank correlation method was used [38].

III. Results

This section consists of three subsections. In the first one, the length of reviews is analyzed between review types. In the second one, the playtimes when writing reviews are examined. The correlation between playtimes before reviewing and review type is also investigated in it. Lastly, in the third subsection, the most frequent words and their associations are analyzed.

A. *Word numbers between review types*

The next step was to compare the reviews grouped by whether they were positive or negative. There were 1,292,758 positive reviews and 343,161 negative ones in the dataset. In both types, there were several reviews that had a length of zero. Still, it can be seen that the negative reviews contained more words than positive ones. The median, mean, and standard deviations of negative reviews were 36; 82.1; and 136.73 words, respectively. On the contrary, their respective words were 18; 53.68; and 107.83 in the case of the positive reviews. The longest negative review was 2,667 words long, while the longest positive one was 2,214 words long. Overall, as can be seen, negative reviews were longer than the positive ones. The distribution of word numbers can be observed in Fig. 1.

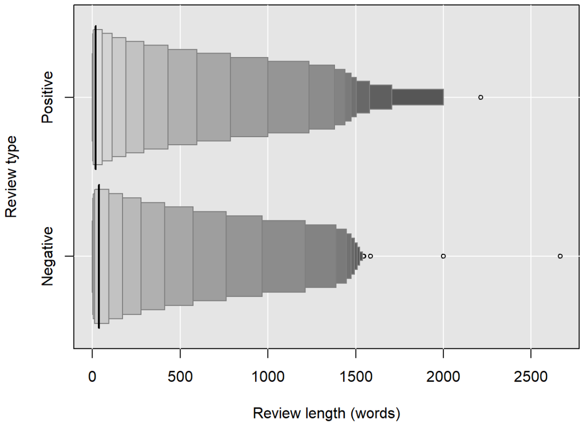


Fig. 1. The distribution of words grouped by review type.

Before using statistical tests to compare the review types to each other, it was also assessed whether word numbers follow Gaussian distribution in the dataset. Due to the large number of data, the previously mentioned Q-Q plots were used for this purpose. As can be seen in Fig. 2, the data did not follow Gaussian distribution.

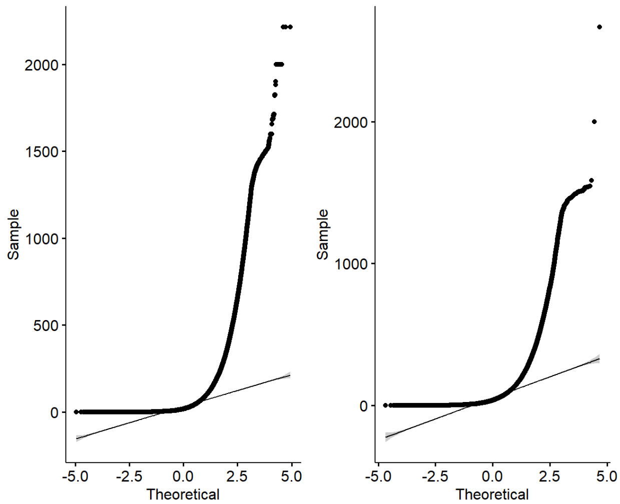


Fig. 2. Q-Q plots of word numbers in positive (left) and negative reviews (right).

Afterward, the Wilcoxon rank sum test was used to see whether the lengths of these review types were significantly different from each other. The results of the test show that the negative ones were significantly longer, $W = 1.6979 \cdot 10^{11}$; $p < 2.2 \cdot 10^{-16}$. It was also assessed with regression analysis methods whether the review type had an effect on review length. The results in Table I show that an effect exists since positive reviews were significantly shorter on average.

TABLE I
RESULTS OF THE REGRESSION ANALYSIS

	Estimate	Standard error	t value	$Pr(>\|t\|)$
Intercept	82.0994	0.1955	419.9	$< 2 \cdot 10^{-16}$
Positive review	-28.4220	0.2200	-129.2	$< 2 \cdot 10^{-16}$

As can be seen in Table I, the average length significantly differs between the two review types. On average, positive reviews contain significantly fewer words than negative ones. The mean difference between the two review types is 28.4220 words.

B. *Playtime between review types*

Next, the playtime of users between the two review types were investigated. According to the results, players review VR games after playing for an average of 2,172 minutes. The median value is 281 minutes, and the standard deviation is 13,502.37 minutes. Clearly, most players tend to play a VR game for several hours before leaving a review for it. It should be noted that the largest playtime before reviewing was 2,867,784 minutes. That would mean 47,796.4 hours. In the case of positive reviews, the mean, median, and standard deviation values are 2278; 335; and 13,478.66 minutes, respectively. Regarding negative reviews, their respective values are 1,773; 99; 13,584.04 minutes. As can be observed, negative reviews are written much earlier than positive ones. The distribution of playtimes between review types can be seen in Fig. 3.

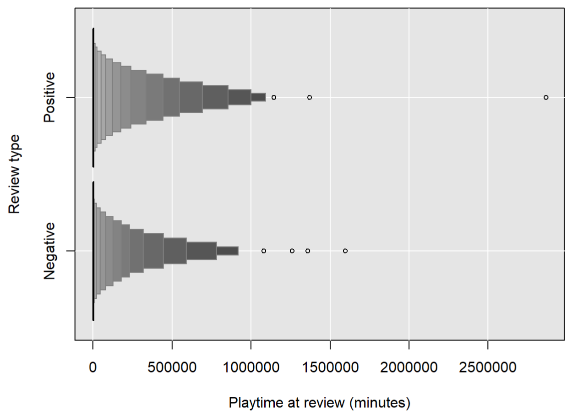


Fig. 3. The distribution of playtime before reviewing grouped by review type.

Similarly, to word numbers, Q-Q plots were used to assess whether playtimes followed Gaussian distribution. These plots can be observed in Fig. 4.

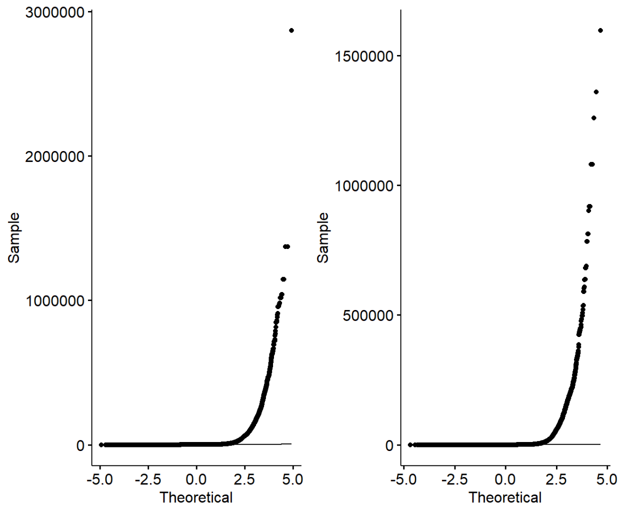


Fig. 4. Q-Q plots of playtimes in positive (left) and negative reviews (right).

As the distribution was non-Gaussian, the Wilcoxon rank sum test was used to compare the playtimes between the two review types. The results show that there is a significant difference between them, $W = 2.8911 \cdot 10^{11}; p < 2.2 \cdot 10^{-16}$. As previously, regression analysis methods were used to see whether review type had a significant effect on playtimes before reviewing. According to the results presented in Table II, a significant effect exists. Positive reviews were written significantly later than negative ones.

TABLE II
RESULTS OF THE REGRESSION ANALYSIS

	Estimate	Standard error	t value	$Pr(>\|t\|)$
Intercept	1,772.74	23.09	76.76	$< 2 \cdot 10^{-16}$
Positive review	505.04	25.98	19.44	$< 2 \cdot 10^{-16}$

It is shown in Table II that level is significance is strong in each case. On average, positive reviews are written significantly later than negative with a difference of 505.04 minutes. This would mean that those who play VR for more hours are more likely to leave a positive review.

However, it is imperative to understand whether a correlation exists between playtime and review length. Therefore, the next step was to assess this. First, this relationship was assessed on the whole dataset. However, no correlation was found, $r_s(1{,}635{,}917) = -0.046;\ p < 2.2 \cdot 10^{-16}$. When only positive reviews were analyzed in this regard, no correlation was found as well, $r_s(1{,}292{,}756) = -0.040;\ p < 2.2 \cdot 10^{-16}$. Then, negative reviews were analyzed, however, they presented a similar relationship between the two variables, $r_s(343{,}159) = 0.106; p < 2.2 \cdot 10^{-16}$. Still, it can be noted that the correlation coefficient was somewhat stronger in this case.

C. The most frequent words and their associations

Next the most frequent words were assessed in both positive and negative reviews. In case of positive ones, they are shown in Figs. 5 and 6.

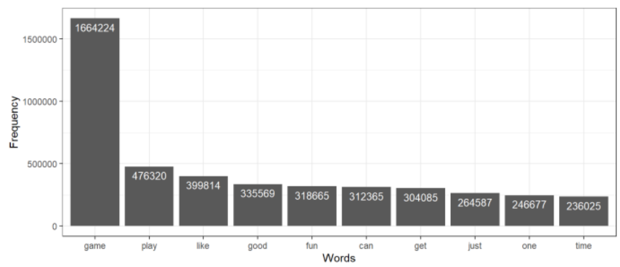


Fig. 5. The most frequent words in positive reviews.

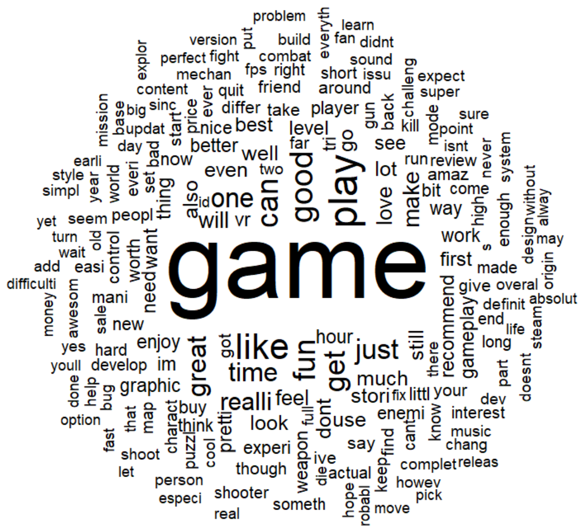


Fig. 6. Word cloud of words in positive reviews.

As can be observed in Figs. 5 and 6, the most frequent word in positive reviews is “game” with 1,664,224 occurrences. The second most frequent was “play” with 476,320 occurrences. Judging from the words, players in positive reviews mention their experiences (e.g. “fun”, “love”, “good”, “great”, “awesome”, “enjoy”), game difficulty, game mechanics, and that they recommend the game. Interestingly, even within positive reviews, there were mentions of bugs, suggesting that while these issues were acknowledged, they did not detract from the overall positive experiences of players. This indicates that the identified bugs may not have severely impacted the overall enjoyment and satisfaction derived from the games, as evidenced by the overwhelmingly positive sentiments expressed by users.

The most frequent words in negative reviews were examined next. They can be observed in Figs. 7. and 8.

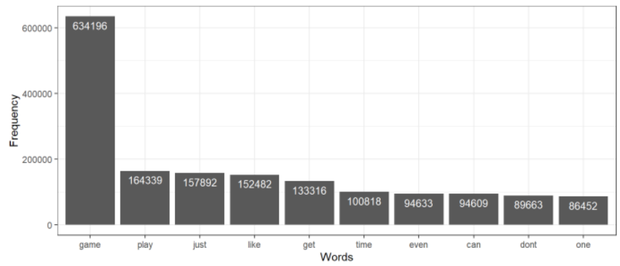


Fig. 7. The most frequent words in negative reviews.

Fig. 8. Word cloud of words in negative reviews.

Similarly to positive reviews, "game" and "play" were the two most frequent words in negative reviews, with 634,196 and 164,339 occurrences, respectively. This highlights the emphasis on gaming experience and active engagement within the VR gaming environment, albeit in a negative context. However, in negative reviews, the word "don't" emerged as a prominent term, replacing "time" in the top 10 most frequent words, indicating dissatisfaction among users. Correspondingly, the negative sentiment was reflected in the increased frequency of words such as "poor", "bad", "bug", "problem", and "crash", signifying a focus on technical problems and performance issues encountered by players. Furthermore, the presence of the word "server" within these reviews suggests that network-related or device-related problems contributed to the negative experiences reported by users. Additionally, the mention of "fix" underscores the imperative for developers to address the identified bugs within the games. Notably, the inclusion of the word "refund" indicates that users expressed a desire for reimbursement due to the unsatisfactory nature of their experiences. Players also showed concerns related to game mechanics, specifically pertaining to combat and weapons, as well as references to gameplay difficulty, as evidenced by the inclusion of the word "hard". Despite many negative sentiments, a few positive words were also present, albeit less frequently. This indicates a nuanced and varied spectrum of user feedback. Moreover, the words in Fig. 8 revealed the emergence of profanity in the reviews, with the f-word being notably prevalent, showing negative emotions and frustration expressed within the negative reviews.

Naturally, these words alone do not give enough context to understand the strengths and weaknesses of VR games. Thus, as was mentioned in Section II, word associations of commonly occurring words were also examined. The results of this investigation can be observed in Fig. 9 in case of both review types.

When examining positive reviews, the following observations could be made. Among others, the words "voice acting" can be commonly found in these reviews. The cases are similar with the words "single player", "trial error", "learning curve", "motion sickness", and "paced fast". Similarly, "extremely", "absolutely", "fantastic", "favorite", "excellent", "totally", "highly" and "recommended" have a strong correlation. Interestingly, profane words are still highly correlated with each other in positive reviews. Even though there were some negative word associations like between "motion" and "sickness", they were not negative enough to warrant a negative review.

Negative reviews however, present different correlations between words. The correlated word with "money" is "waste", indicating that players regretted buying the game. Similarly, "alt" correlates with "f4", suggesting that players frequently quit the games due to anger or frustration. The word

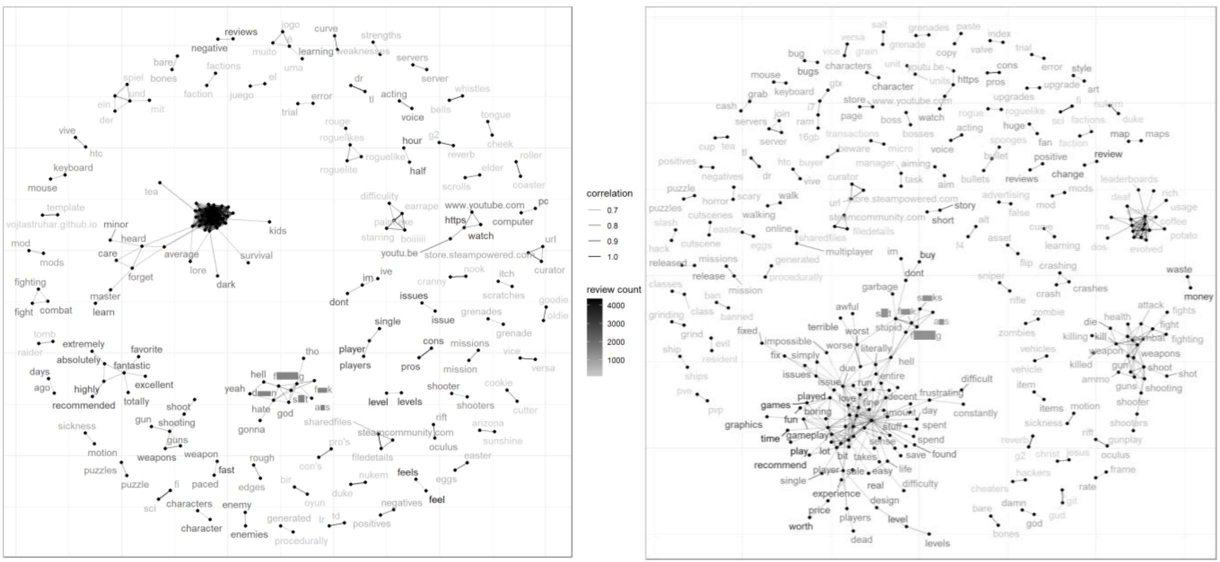


Fig. 9. Frequent words and their associations in positive (left) and negative reviews (right).

"advertising" correlates with "false", showing that even the ads for these games contained false information. "Short" is commonly found associated with "story", indicating that the length of these games was also a problem. Looking at the second largest cluster of word correlations on the right, it can observed that game mechanics regarding shooting with guns, or simply battle mechanics were badly designed, resulting in in-game character death. When examining the largest cluster of word correlations, the following can be stated. Firstly, those words appeared in most negative reviews. Secondly, the cluster of words is full of profanity and negative adjectives. Thirdly, these words can be associated with the game's difficulty or boring gameplay, bad level design, bugs/issues, and not recommending the games. Interestingly, "sale" is also mentioned, indicating that the games should be bought during a sale as they were not worth their full asking price.

IV. Discussion

The research questions were answered by the findings which provided an understanding of VR game reviews. Thus, in this section, the implications of the results are discussed.

The first research question focused on understanding the length of VR game reviews. The results indicate that negative reviews tend to be longer on average compared to positive ones. This suggests that players express their dissatisfaction in greater detail. This is similar to the results presented by Lin et al. [29]. In their study, they compared the length of reviews between various video game genres. Although they did not assess VR games, our results show that a similar phenomenon exists in their case.

The second research question revolved around the relationship between playtime and word numbers in the reviews. While no correlation was found, there was a significant difference in when players wrote the reviews. Positive reviews are written much later than negative ones. This raises interesting questions about player motivations. Negative reviews are written earlier due to frustration or disappointment, while positive ones are written after a more extensive gameplay session that strengthens positive impressions. These results are also similar to the ones concluded by Lin et al. when they examined review times between video game genres [29]. Our results show that VR games show a similar phenomenon in this regard.

The third research question focused on the understanding of word frequencies and their associations in the reviews. The results provided valuable contextual information about player experiences. Positive reviews highlight the players' appreciation for narrative elements, gameplay, and game mechanics. On the contrary, negative reviews often emphasized dissatisfaction with purchasing decisions and they showed the importance of how players evaluate the value of money. Similarly, players were critical of game length and narrative depth. The strong correlations between certain words, particularly profanity and negative adjectives, highlights the emotional and critical nature of negative reviews. These findings show that players use specific language to express their frustrations with game mechanics, and certain design elements.

This study holds significant implications for both VR game developers and researchers. For developers, gaining insights into the correlation between review length and playtime can serve as a valuable foundation for creating new strategies that are aimed at enhancing player engagement and immersion. By addressing the highlighted concerns within negative reviews, such as game mechanics, level design, and fixing the identified bugs, developers can actively contribute to an improved player satisfaction and experience. Moreover, the word associations within the reviews present an opportunity to see the preferences and priorities of players. With this understanding, developers can optimize and refine aspects of their games that resonate positively with players, while simultaneously identifying and solving issues that create negative sentiments. Furthermore, the identification of prevalent terms within positive reviews, a roadmap can be provided for developers. With it, they can reinforce the strengths of their games and bolster player satisfaction as well. Using these results, developers can more effectively tailor their future projects to align more closely with the preferences and expectations of their player base.

Naturally, this study has its limitations. Firstly, only English reviews were scraped, and the limit of reviews was 1,000 per VR game. Secondly, the analysis was restricted to textual content. Thirdly, only Steam reviews were investigated. Future research could expand the analysis by including other digital video game distribution platforms and other sources of feedback, such as forum discussions or social media conversations. Also, sentiment analysis techniques and natural language processing algorithms could be included to provide a deeper understanding of emotions in the reviews.

V. Conclusions

In this study, we analyzed VR game reviews to gain a deeper understanding of player experience. We investigated 1,635,919 textual reviews comprising both positive and negative reviews from the Steam digital video game distribution platform. Our goal was to understand review length, playtime before reviewing, and frequent words as well as their associations regarding VR games.

The results within this study show that negative reviews are written significantly earlier and contain more words than positive ones. This suggests that players tend to write about their dissatisfaction as early as possible and they express their concerns in greater depth. On the contrary, positive reviews contain concise language. These positive reviews also highlight the significance of narrative elements, gameplay, and learning experiences. Negative reviews, on the other hand, range from monetary concerns to gameplay mechanics, and even false advertising. The resulting word correlations provided a glimpse into players' experience.

In conclusion, our results can provide recommendations for developers and researchers alike. By understanding the relationships between review length, playtime, and word frequency as well as their associations, richer player experiences can be created in the realm of VR.

Acknowledgment

This work has been implemented by the TKP2021-NVA-10 project with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021 Thematic Excellence Programme funding scheme.

References

[1] G. Wersényi, "Musings on the Metaverse and on how the Internet of Digital Reality encompasses future developments of the Internet," *Infocommunications Journal*, vol. 15, no. Special Issue: Internet of Digital & Cognitive realities, pp. 21–28, 2023, DOI: 10.36244/ICJ.2023.SI-IODCR.4.

[2] F. Erdős, R. Németh, and B. Firuza, "Virtual Teamwork in Gamified 3D Environment," *Infocommunications Journal*, vol. 15, no. Special Issue: Internet of Digital & Cognitive realities, pp. 15–20, 2023. DOI: 10.36244/ICJ.2023.SI-IODCR.3.

[3] F. Faiqotuzzulfa and S. A. Putra, "Virtual reality's impacts on learning results in 5.0 education: A meta-analysis," *International Transactions on Education Technology (ITEE)*, vol. 1, no. 1, pp. 10–18, 2023, DOI: 10.33050/itee.v1i1.172.

[4] S. G. Izard, J. A. Juanes Méndez, and P. R. Palomera, "Virtual reality educational tool for human anatomy," *Journal of Medical Systems*, vol. 41, no. 5, 2017, DOI: 10.1007/s10916-017-0723-6.

[5] M. A. Rojas-Sánchez, P. R. Palos-Sánchez, and J. A. Folgado-Fernández, "Systematic literature review and bibliometric analysis on virtual reality and education," *Education and Information Technologies*, vol. 28, no. 1, pp. 155– 192, 2023, DOI: 10.1007/s10639-022-11167-5.

[6] P. Smutny, "Learning with virtual reality: a market analysis of educational and training applications," *Interactive Learning Environments*, pp. 1–14, 2022, DOI: 10.1080/10494820.2022.2028856.

[7] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F. D. Lobo, C. Li, H. Huang, M. Akdere, C. Mousas, and L.-F. Yu "A review on virtual reality skill training applications," *Frontiers in Virtual Reality*, vol. 2, 2021, DOI: 10.3389/frvir.2021.645153.

[8] D. W. Carruth, "Virtual reality for education and workforce training," in *2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, Stary Smokovec, Slovakia, 2017, pp. 1–6, DOI: 10.1109/ICETA.2017.8102472.

[9] N. Gasteiger, S. N. van der Veer, P. Wilson, and D. Dowding, "How, for whom, and in which contexts or conditions augmented and virtual reality training works in upskilling health care workers: Realist synthesis," *JMIR Serious Games*, vol. 10, no. 1, p. e31644, 2022, DOI: 10.2196/31644.

[10] M. Matsangidou, F. Frangoudes, M. Hadjiaros, E. Schiza, K. C., Neokleous, E. Papayianni, M. Avraamides, and C. S. Pattichis "'Bring me sunshine, bring me (physical) strength': The case of dementia. Designing and implementing a virtual reality system for physical training during the COVID-19 pandemic," *International Journal of Human-Computer Studies*, vol. 165, no. 102840, p. 102840, 2022, DOI: 10.1016/j.ijhcs.2022.102840.

[11] S. Z. A. Ansari, V . K. Shukla, K. Saxena, and B. Filomeno, "Implementing virtual reality in entertainment industry," in *Cyber Intelligence and Information Retrieval*, Singapore: Springer Singapore, 2022, pp. 561–570, DOI: 10.1007/978-981-16-4284-5_49.

[12] F. Rega and D. Saxena, "Free-roam virtual reality: A new avenue for gaming," in *Advances in Augmented Reality and Virtual Reality*, Singapore: Springer Singapore, 2022, pp. 29–34, DOI: 10.1007/978-981-16-7220-0_3.

[13] G. Burdea and P. Coiffet, *Virtual Reality Technology*. Nashville, TN: John Wiley & Sons, 1994.

[14] I. Horváth, Á. B. Csapó, B. Berki, A. Sudár, and P. Baranyi, "Definition, background and research perspectives behind 'Cognitive Aspects of Virtual Reality' (cVR)," *Infocommunications Journal*, vol. 15, Special Issue: Internet of Digital & Cognitive realities, pp. 9–14, 2023, DOI: 10.36244/ICJ.2023.SI-IODCR.2.

[15] C. V ogel and A. Esposito, "Interaction Analysis and Cognitive Infocommunications," *Infocommunications Journal*, vol. 12, no. 1, pp. 2–9, 2020, DOI: 10.36244/ICJ.2020.1.1.

[16] H. Mclellan, "Cognitive issues in virtual reality," *Journal of Visual literacy*, vol. 18, no. 2, pp. 175–199, 1998, DOI: 10.1080/23796529.1998.11674538.

[17] S. Hedegaard and J. G. Simonsen, "Extracting usability and user experience information from online user reviews," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, DOI: 10.1145/2470654.2481286.

[18] H.-N. Kang, "A study of analyzing on online game reviews using a data mining approach: STEAM community data," *International Journal of Innovation, Management and Technology*, pp. 90–94, 2017, DOI: 10.18178/ijimt.2017.8.2.709

[19] C. Chambers, W.-C. Feng, S. Sahu, and D. Saha, "Measurement-based characterization of a collection of on-line games," in *Proceedings of the 5th ACM SIGCOMM conference on Internet measurement – IMC '05*, 2005.

[20] I. Busurkina, V. Karpenko, E. Tulubenskaya, and D. Bulygin, "Game experience evaluation. A study of game reviews on the steam platform," in *Communications in Computer and Information Science*, Cham: Springer International Publishing, 2020, pp. 117–127, DOI: 10.1007/978-3-030-65218-0_9.

[21] X. Li, Z. Zhang, and K. Stefanidis, "A data-driven approach for video game playability analysis based on players' reviews," *Information (Basel)*, vol. 12, no. 3, p. 129, 2021, DOI: 10.3390/info12030129.

[22] S. Ahn, J. Kang, and S. Park, "What makes the difference between popular games and unpopular games? analysis of online game reviews from steam platform using word2vec and bass model," *ICIC Express Letters*, vol. 11, pp. 1729–1737, 2017.

[23] K. Wase, P. Ramteke, R. Bandewar, N. Badole, B. Kumar, and M. Bogawar, "Sentiment Analysis of product review," *International Journal of Innovations in Engineering and Science*, vol. 3, pp. 8–13, 2018.

[24] G. Andreev, D. Saxena, and J. K. Verma, "Impact of review sentiment and magnitude on customers' recommendations for video games," in *2021 International Conference on Computational Performance Evaluation (ComPE)*, Shillong, India, 2021, pp. 992–995, DOI: 10.1109/ComPE53109.2021.9752380.

[25] V. de Souza Goncalves and A. A. Macedo, "Similarities and divergences in electronic game review texts," in *2020 19th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, Recife, Brazil, 2020, pp. 145–153, DOI: 10.1109/SBGames51465.2020.00028.

[26] E. Guzman and W. Maalej, "How do users like this feature? A fine grained sentiment analysis of app reviews," in *2014 IEEE 22nd International Requirements Engineering Conference (RE)*, Karlskrona, Sweden, 2014, pp. 153–162, DOI: 10.1109/RE.2014.6912257.

[27] C. Lu, X. Li, T. Nummenmaa, Z. Zhang, and J. Peltonen, "Patches and Player Community Perceptions: Analysis of No Man's Sky Steam Reviews," in *DiGRA'20-Proceedings of the 2020 DiGRA International Conference*, 2020, pp. 1–24.

[28] Y. Yu, B.-H. Nguyen, F. Yu, and V.-N. Huynh, "Esports game updates and player perception: Data analysis of PUBG steam reviews," in *2021 13th International Conference on Knowledge and Systems Engineering (KSE)*, Bangkok, Thailand, 2021, pp. 1–6, DOI: 10.1109/KSE53942.2021.9648670.

[29] D. Lin, C.-P. Bezemer, Y. Zou, and A. E. Hassan, "An empirical study of game reviews on the Steam platform," *Empirical Software Engineering*, vol. 24, no. 1, pp. 170–207, 2019. DOI: 10.1007/s10664-018-9627-4.

[30] B. Dean, "Steam usage and catalog stats for 2023," *Backlinko*, 17-Mar-2023. [Online]. Available: https://backlinko.com/steam-users. [Accessed: 19-Aug-2023].

[31] E. Blancaflor and J. M. G. San Miguel, "Analyzing Digital Game Distribution in Gaming Industry: A Case Study," in *Proceedings of the 2nd Indian International Conference on Industrial Engineering and Operations Management*, 2022, pp. 674–681.

[32] Valve Corporation, “Documentation home page (steamworks documentation),” *Steamgames.com*. [Online]. Available: https://partner.steamgames.com/doc/home. [Accessed: 19-Aug-2023].
[33] Valve Corporation, “User reviews – get list (steamworks documentation),” *Steamgames.com*. [Online]. Available: https://partner.steamgames.com/doc/store/getreviews. [Accessed: 19-Aug- 2023].
[34] Valve Corportation, “ISteamApps interface (steamworks documentation),” *Steamgames.com*. [Online]. Available: https://partner.steamgames.com/doc/webapi/ISteamApps. [Accessed: 19-Aug-2023].
[35] C. J. Van Rijsbergen, S. E. Robertson, and M. F. Porter, “New models in probabilistic information retrieval,” in *British Library. (British Library Research and Development Report*, London, 1980.
[36] H. Hofmann, H. Wickham, and K. Kafadar, “Letter-value plots: Boxplots for large data,” *Journal of Computational and Graphical Statistics*, vol. 26, no. 3, pp. 469–477, 2017,
DOI: 10.1080/10618600.2017.1305277.
[37] H. B. Mann and D. R. Whitney, “On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other,” *The annals of mathematical statistics*, vol. 18, no. 1, pp. 50–60, 1947.
[38] C. Spearman, “The proof and measurement of association between two things,” *American Journal of Psychology*, vol. 15, no. 1, p. 72, 1904, DOI: 10.2307/1412159.

Tibor Guzsvinecz is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. He obtained his PhD degree in the field of Information Science in 2021. His research areas are data science, gamification, human- computer interaction, spatial ability, and virtual reality. He authored 69 scientific publications which include 19 journal papers, and 41 studies found in various conference proceedings. At the moment, he is an editor of a book project about gamification. He is an IEEE member since 2023.

Judit Szűcs is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. She received her PhD degree in 2021 in Computer Science at the University of Szeged. Her main interests include binary tomography, image processing, and computer vision. She has 30 scientific publications which include 6 journal papers, and 19 studies found in various conference proceedings. She is an IEEE member since 2023.

Special Issue
of the Infocommunication Journal

Design and Evaluation of Abstract Aggregated Avatars in VR Workspaces

György Persa*, and Ádám B. Csapó†‡

Abstract—**Avatars are commonly used in digital platforms to provide a visual representation of individual users to each other. Generally, avatar design in the past has focused on achieving visual fidelity and realism of social interactions. In this paper, we broaden the concept of avatars to incorporate displays using an abstract visual language and conveying information on aggregated, interpersonal information from the perspective of the digital platform as a whole. We propose a general design methodology for such aggregated avatars, and also introduce and experimentally evaluate an aggregated avatar which we have developed on the MaxWhere VR platform. Results are promising in that users were able to discern several key states of the avatar and correctly associate them with the correct virtual reality scenarios in a statistically meaningful way.**

Index Terms—**virtual reality, avatar design, abstract avatars, aggregated avatars**

I. Introduction

The word '*avatar*' originates from Sanskrit, and refers to the meaning of 'embodiment', or a 'divine being made of flesh' [1]. In the rapidly evolving fields of virtual reality and metaverse technology, it evokes a very specific connotation of a human-like visual representation that can convey real-time information on the appearance, activities and even the mental state of a given user [2], [3].

Despite this clear picture of what an avatar is, or should be, it is nevertheless worth noting that avatars can take many shapes and forms, depending on the specific information they are used to represent, the visual and cognitive fidelity of their representation, and even depending on the context in which they are used. At one 'corner' of this multi-dimensional spectrum, an avatar can appear as a simple pictogram, showing a static photo of the user and indicating their presence on the social / virtual platform. At another 'corner' of the spectrum, highly visually realistic avatars are used that digitally re-represent the user's facial and head movements based on real-time tracking. In a third 'corner' of the spectrum, one can consider digital environments that involve many users interacting at once, instead of focusing on one-on-one, or small-group social interactions. In the case of a personalized virtual reality working environment, helping users to either streamline their digital workflows, or to focus on personal productivity and personal growth in an engaging environment would be the most relevant feature. In such scenarios, avatars that provide motivation to users by reflecting back to them key insights on their activities as a kind of meta-cognitive feature of the platform might be most valuable. Such meta-cognitive functionality for an avatar can also in fact be considered in cases where a platform has many users and the aggregate statistics reflecting the way in which the platform is being used will be of more interest than the interaction patterns of any given user. In this scenario, each user might contribute in a small way to the overall behavior of the avatar.

In this paper, we broaden the traditional understanding of avatars to include representations that mirror not just the specifics of a single user to others but also provide a holistic view of the digital environment to the user. This includes the state of the digital platform and the interactions of the user and possibly numerous other users with the platform. By easing the rigid definitions of what an avatar represents and to whom, we propose the possibility of previously unexplored avatar types, which we label as 'abstract avatars' and 'aggregated avatars'. We further posit that these types of avatar hold significant potential in our evolving digital landscape, where human and digital interactions are deepening not only in the short run but also in a more sustained, co-evolutionary manner as often described in the literature on cognitive infocommunications and cognitive aspects of virtual reality [4], [5]. To validate this concept, we describe a possible design methodology for such avatars, develop an example implementation, and perform validation of this implementation to demonstrate the viability of this approach.

The paper is structured as follows. In Section II, we consider different dimensions along which avatars can be qualified, and define novel categories of avatars based on these dimensions, including abstract and aggregated avatars. In Section III, we describe the design steps we have taken to design an aggregated avatar based on emotional features. In Section IV, we describe a framework within which we validate the design of our aggregated avatar through test subjects. Finally, Sections V and VI present details on the experimental design we have used, and the results we have obtained.

II. Conceptualizing novel types of avatars

A. Visual and cognitive fidelity

As discussed in the previous section, the visual representations used in avatars can range from simple, static pictograms to highly accurate visual representations of specific users [2], [6]–[10]. Based on this, it is possible to consider the *visual*

* Doctoral School of Multidisciplinary Engineering Sciences, Széchenyi István University, Győr Hungary
† Corvinus Institute for Advanced Studies & Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Hungary
‡ Hungarian Research Network, Budapest, Hungary
E-mail: persa.gyorgy@sze.hu
E-mail: adambalazs.csapo@uni-corvinus.hu

DOI: 10.36244/ICJ.2024.5.11

fidelity of an avatar, in terms of how it relates to the user it represents in a visual sense.

At the same time, avatars can also be assessed in terms of their behaviors, whether in terms of how naturally they behave in social interactions [11]–[15], or in terms of how well they reflect the key aspects of user interactions, thereby contributing to increased productivity. Based on this, we have defined the term *cognitive fidelity* as follows [16]:

Definition 1: The **cognitive fidelity** of an avatar is a an assessment, both qualitative and potentially quantitative, of how well an avatar reflects the state of an underlying process, viewed from the attendant benefits to the cognitive capabilities of the users to whom it appears.

We note that the term cognitive fidelity has been used in other contexts, for example, to describe how well the use of a virtual tool corresponds to users' actions and possible choices [17], [18]; or in other cases, to describe the ecological validity of an environment from the standpoint of cognitive tasks being carried out inside it [19]. In the above definition, we focus instead on the mental capabilities of the user to whom the avatar is displayed.

As an example, let us consider a flight simulator. In such an application, the designer would have a choice as to whether to represent the co-pilot of the user as a photo-realistic human whose body posture and head movements are fully in accordance with normal social interactions, or instead, to use more abstract representations that can offer useful feedback as to what is going wrong or what aspects of flight control require attention. In some cases, even the feedback that "something is going wrong" can be highly relevant and can improve cognitive performance.

B. Abstract and Aggregated Avatars

When the main focus is on cognitive rather than visual fidelity, questions of information modeling and representation mapping naturally arise.

In terms of information modeling, it is often the case that parameters that are not so directly linked with physical reality (for example, statistics on different interaction types, or other parameters closely related to the application scenario) need to be communicated through the avatar.

In terms of finding a useful representation, the goal is to map the information to be represented onto visual (and perhaps other sensory) channels in a way that is recognizable and intuitively meaningful to users. From the theory of cognitive infocommunication (CogInfoCom) channels, if we consider avatars to be analogous to CogInfoCom messages, then the information mapping types of direct (both low-level and high-level), as well as indirect (including structural, co-stimulation or scenario-based) can be especially relevant [4], [20].

Based on these considerations, we have introduced the following definitions of abstract and aggregated avatars, with the intention of encompassing a broader scope of potential avatar designs [16]. Here, we propose a slightly revised version of this definition so as to focus – in the case of abstract avatars – solely on the language of the avatar, irrespective of the application scenario:

Definition 2: An **abstract avatar** is an avatar representation that relies on an abstract, low-level visual language involving the use of dynamic shapes and colors without reference to anthropomorphic or zoomorphic concepts.

Definition 3: An **aggregated avatar** is an abstract avatar that is used to display the features of an impersonal set of interactions and contextual events in a computational environment.

On the one hand, aggregated avatars are capable of providing users with a mirror of their position within a cooperative process, as seen through the perspective of the digital environment, instead of merely presenting data about other users. Conversely, due to its largely impersonal character, aggregated avatars can also serve to inform users about the comprehensive 'status' of a platform like a workplace or a virtual reality setting. This includes collective, environmental concepts such as the 'vibrancy of the surroundings' or the 'enthusiasm level of participants'. These aspects are somewhat influenced by the conduct of the users involved, yet they cannot easily be broken down into the distinct individual behaviors of each user.

C. Design principles for the development of abstract and aggregated avatars

Based on the dichotomy of information and representation modeling, the design of an abstract or aggregated avatar can be broken into the following steps:

1) Answering the question of what information types to model, including potential range of numerical values. In case the state to be modeled includes (fuzzy) categorical features as well, e.g. fuzzy modeling can be used to convert states into numerical fuzzy membership values.
2) Answering the question of how the resulting variables can be mapped onto an intermediate language suitable for driving discernible and interpretable avatar behaviors. One example of such an intermediate language can be a valence-arousal based emotional model, which is capable of representing distinct emotional states that can be recognized with relative ease by many users.
3) Finally, defining the visual features of the abstract avatar that can be used to drive its behavior, and mapping onto them the states of the intermediate language described in the previous point.

III. Design of an Emotionally Based Aggregated Avatar

We have given design principles for aggregated avatars in previous sections, aiming to enhance cognitive fidelity by finding suitable mappings between cognitive attributes of users and visual features of avatars. In our prior works, we created and implemented an abstract avatar inspired by ethology that can express emotional states [21]. We used this avatar design to present the aggregated space data to users.

A. Emotion displaying agent

The abstract agent that we designed to display emotions has two distinct parts: A colored sphere that can change

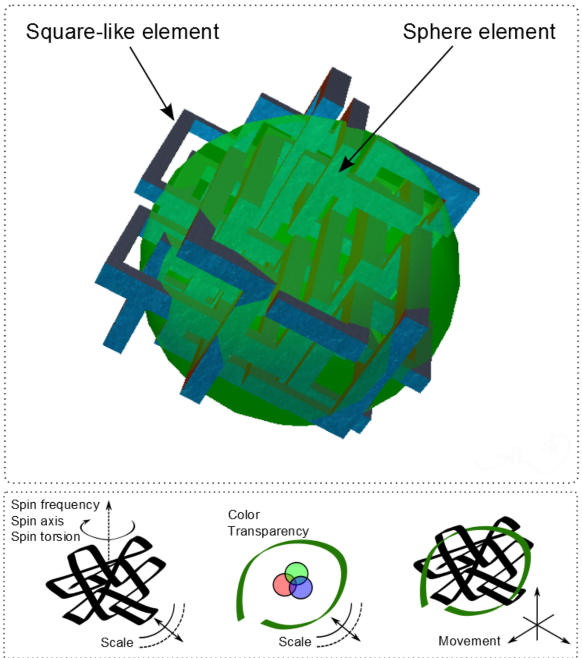


Fig. 1. The appearance and degrees of freedom of the emotional display agent

its transparency and size; and a cube-like structure that has different colors on each side and resembles a maze 1. Degrees of freedom included the position of the component, size of elements separately, color and transparency of the sphere element and rotation speed and angle of the square-like element. This setup provided us with the possibility of evoking associations with emotional states [22].

A possible way to illustrate the aggregated state of a virtual environment is to use the emotion displaying agent as a collective representation of the space. The agent acts as an abstract aggregated avatar that reflects the overall mood of the virtual space through its visual features. This approach requires a mapping function that can translate the aggregated state of the space into the visual parameters of the agent's appearance and behavior.

B. Mapping emotions to agent

Based on the evaluation of expressive features of the Emotion Display Agent, we designed a continuous mapping which can construct the look of the Emotion Display Agent for any combination of emotions. For this purpose we used the well-known Valence-Arousal model of emotions. This psychological framework describes how emotions are represented in two dimensions: valence refers to the positive or negative quality of an emotion, while arousal refers to its intensity or activation level.

The main goal with the mapping was to produce the kinds of appearances in the case of different valence-arousal value compbinations that have already been associated with corresponding emotion (happiness, sadness, fear, etc.) by users in the previously cited study [22]. Thus, we designed the following rules to achieve this:

- High valence and high arousal drove the avatar into a state where it expressed happiness by growing in size, with the sphere element exceeding the inner maze shape and appearing in a yellowish color while rotating at a relaxed pace
- High arousal with low valence caused the avatar to display anger: the sphere element turned red and shrunk while the maze element exceeded the sphere in size and rotated with high speed and non-linear easing.
- Low valence with low arousal caused the avatar to display a sad expression. It rotated evenly, moved away from the camera and shrunk in every dimension while the maze elements greatly exceeded the size of the sphere. The color looked approximately pale purple in this state
- High valence with low arousal reflected a bored or sleepy state of the avatar. To express this emotion, the maze element rotated at a low speed but unevenly, while the sphere grew big and appeared in a purple-like color
- In its natural state the avatar looked relaxed, with a green colored sphere slightly exceeding the size of the maze element. Rotation and easing were adjusted to normal pace in this case.

The top-right corner of Figure 2 shows several examples of how arousal and valence were mapped onto the avatar parameters.

C. Mapping aggregated interactions to emotions

As described in [5], modern infocommunication platforms are evolving to encompass a wide variety of novel interfaces, including virtual reality interfaces, AI-driven interfaces and distributed Web 3.0 applications. In this context, new kinds of aggregated information types are emerging which, when communicated to users in a way that represents the system as a whole, can provide intuitive feedback on parameters such as how active, how overloaded, or how quiet the system is – whatever the case may be.

In the context of virtual reality, parameters such as number of users, activity level of users in terms of moving around in the space, exploring localized subsets of the space, or interacting with others in the space could be of interest.

The specific mapping functions that are used will naturally depend on the aggregated data and the goal of the system in communicating it to users. In Section IV-B, we present a detailed example of one specific application we have developed to test the aggregated avatar concept.

IV. A Test Framework for Evaluating Abstract and Aggregated Avatars

To evaluate the effectiveness of the abstract aggregated avatar, we developed a test framework in the MaxWhere VR application. The main objective was to obtain quantitative data

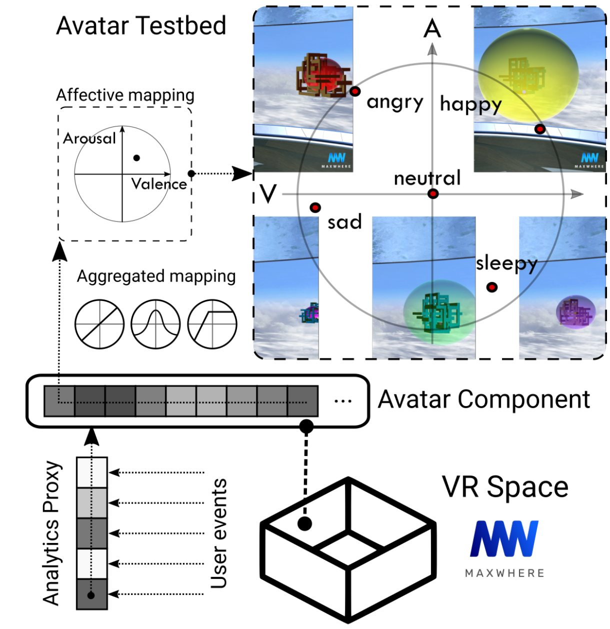


Fig. 2. Architectural diagram of testbed framework used for investigating aggregated avatar in MaxWhere VR.

on how well the participants could understand and differentiate the aggregated state of the virtual space. The basic idea was to design a test in which we could present various user interaction scenarios in the virtual environment and also display the state of the abstract aggregated avatar. We could then test whether a participant could identify the relationship between a user interaction and the corresponding changes in the visual appearance of the avatar by either swapping or keeping aligned the interaction videos and the avatar recordings, and repeating this idea with several different kinds of interactions. The overall structure of the framework is shown in Figure 2.

A. MaxWhere VR

MaxWhere VR is an innovative virtual reality platform that enables users to create and experience immersive virtual environments for various applications. MaxWhere VR offers a versatile and customizable framework that can suit different needs and goals. Users can access a rich library of 3D scenes or import their own content to design their own virtual worlds with custom functionalities. It can be used for presenting, showcasing, collaborating or research purposes as well.

Unlike other platforms, MaxWhere VR offers a unique feature called the “Where Object Model” (WOM), which is similar to a Document Object Model (DOM) used in web programming, but for 3D spaces. The WOM allows users to access and manipulate the properties and behaviors of 3D objects using Javascript code. Users can also create and load components, which are reusable pieces of code that can enhance the functionality and interactivity of the 3D spaces. With MaxWhere, users can design and program their own 3D spaces in a flexible and dynamic way, using powerful programming interfaces.

The platform also provides users with so-called “smartboards”, which are essentially 2D display panels that can be used to show web content in a customizable size, position and orientation within the 3D space.

Due to its ability to customize the environment MaxWhere VR is convenient for designing scientific experiments in virtual spaces. Adding custom functionality to virtual spaces allows researchers to perform an experiment directly in the virtual space or collect data from the application for further analysis. The business logic of the experiment can be implemented as a component and can be easily added to any MaxWhere space. Many examples of such environments can be found in the literature (see e.g. [20], [23]–[25]).

B. Analytics Proxy and Testbed Component

MaxWhere provides several proxies to aid component developers accessing higher level information about 3D objects or application state. Proxies are modules which use the MaxWhere engine WOM API to implement functionalities and encapsulate them into a dedicated interface. Interfaces defined this way can be made available for MaxWhere component developers by exposing functions to the WOM API thus making them available for any MaxWhere user. Using this structural possibility we developed a Space Analytics MaxWhere Proxy with the aim of reporting about the user interaction in the space conveniently. The proxy provides basic information about the virtual environment, such as the sum axis aligned bounding box (AABB) of the space calculated from each visual object located in the space. Furthermore event listeners can be registered via the interface which reports about the following events:

- User (camera) movement in the space
- User changes between 3D and menu of the space
- User discovers a 3D object by orbiting around it
- User interact with a smartboard in the space with mouse cursor

The reporting methods for these events can be adjusted on the proxy. Users can specify the reporting interval if the continuous event sending is not desirable. Also, the behavior for unchanged states can be modified as the user can choose whether the idle system should report the identical events or not.

Using the generic Space Analytics Proxy we implemented a custom Testbed MaxWhere component for creating statistics from the space data. For the calculations we used a customizable time window (default is 40 seconds) in which each type of data is aggregated. We construct the following properties from the gathered information:

- Ratio of discovered area of the space relative to the total AABB. Covered area is calculated by creating an AABB from camera positions
- User mobility: Ratio of time the user spent moving
- Ratio of time the user spent interacting on a smartboard. Calculated from mouse cursor moves while 2D content is displayed
- Ratio of time the user spent in a smartboard, on the menu or in 3D

- Ratio of time the user spent orbiting around a 3D object

The produced statistics are then converted into the intermediate interpretation of Arousal and Valence values. For this calculation we defined the processing functions for each parameter of the aggregated Space Analytics data. We used three different transition functions: One for a high effect with a quick slope, one which increases the effect of a parameter as it nears the average value and one for a small contribution with stretched out slope. Using these transfer functions we tailored the Arousal and Valence values so that:

- High valence was associated with the statistical parameter values all being around average, which means that the user interactions are not one-sided and the space capabilities are all used for some degree.
- High Arousal values are directly but not linearly correlated with user movements and orbiting around objects, or with the discovered area increasing. Switching between 3D space and 2D menu / interaction on smartboards only had a limited effect of this value.

The calculated affective values were then used to drive the look of the aggregated avatar inside the MaxWhere space. The avatar – as summarized in section III-A - is able to express emotions via changes in its visual parameters. To map affective values derived from spatial events to the adjustable attributes of the abstract avatar we use the mappings presented in section III-B. The data flow behind the framework is presented on Figure 2.

Note that we have also implemented utilities for recording and replaying user interactions and avatar states. When recording, the avatar is temporary hidden and the user can interact with the space normally. Each interaction received on the Analytics Proxy is recorded and written into a file. The proxy is configured to report every state of the space periodically, thus a continuous sequence of states is stored as a result. Using the replay function, it is then possible to read these files at a later time and to drive the aggregated avatar with the recorded space state values. During replay the space is hidden and only the avatar is visible. Hiding the avatar or the space in these utilities was important for creating sample sequences and videos for further analysis.

V. Experimental design

In the following section we describe the details of the experiment we performed to validate the feasibility of our aggregated abstract avatar.

A. Test videos

We used the record features of the MaxWhere testbed component to create several pairs of test videos to evaluate the performance of the avatar. Each pair of videos included a user interaction video and an avatar behavior video. The user interaction videos captured only the scene and the interactions without the avatar, while the reactions of the avatar to these interactions were recorded from the saved log file using the replay function. During replay, we recorded the corresponding video showing only the expressions of the avatar. To facilitate comparison, we selected three different aspects of user behavior as dimensions for choosing typical interaction patterns for the videos. The three dimensions were as follows.

- *Discovery*: Spatial behavior when the user moves around the virtual space and discover multiple 3D objects. It consists of the space analytics data of camera movements and orbiting.
- *Manipulation*: Describes the scenario when the user interacts with the 3D objects using the Editor capabilities of MaxWhere. It includes a little bit of orbiting from the space analytics data and mostly specified by menu transitions and cursor movements. During this scenario the user opens and closes the menu multiple times, inputs values on the user interface or moving around 3D object or resize them in the space with dragging gesture of the cursor.
- *2D Operation*: Characterized by smartboard interactions. Using the space analytics data of entering or leaving a smartboard and cursor moments while a smartboard is selected for usage. This is the case of browsing the web or working on a content in smartboards.

We recorded 8 pairs of videos in total along these dimensions. For each interaction types we made two videos where only that behavior was shown (six videos in total). We also made one video where the user showed no behavior and one video where the user showed all the behaviors. The original videos were about one minute long, but we sped them up to 30 seconds for the experiment.

A frame from a video of a user interaction and the corresponding avatar behavior can be seen on Figure 3.

B. Survey

We conducted an experiment to test how well participants could match user interaction videos with avatar behavior videos. We have chosen the aggregated state of the virtual space to be the *"usefulness"* of user interactions. This means that the "mood" of the affective avatar was driven by the user interaction types and based on the quality and the quantity of interactions the avatar displayed a combination of "happy", "frustrated" or "angry" state.

We created 8 video pairs in total, each consisting of a user interaction video and a corresponding avatar behavior video as described in the previous section. We then randomly selected 2 video pairs for each task and presented them to the participants in a mixed order: first a user interaction video, then another user interaction video, then an avatar behavior video, and then another avatar behavior video. The participants had to decide which of the two avatar behavior videos matched the first user interaction video they saw. In other words, they had to choose between the combinations of (1-3; 2-4) and (1-4; 2-3). This task is presented to the participants four times with different video pairs each time.

The complete experiment consisted of the following steps:

- Language selection (we prepared an English and a Hungarian version)
- First we introduced the details of the research and explained the tasks ahead

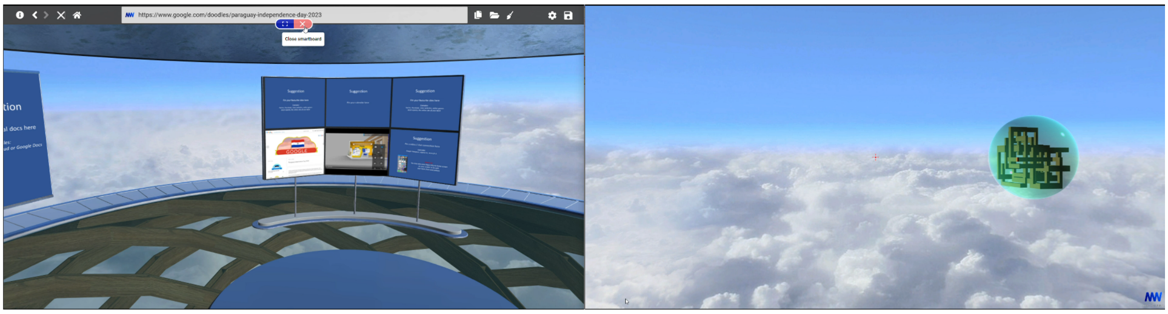


Fig. 3. Example of user interaction video (on the left) with the corresponding avatar video (on the right) showing a 2D operation scenario.

- Collection of informed consent from the participants for anonymized data collection
- Presentation of the 4 matching tasks

The briefing text was the following:

> "*This study explores the expressive capabilities of an abstract-shaped avatar.*
> *The avatar aims to convey a summary (aggregated state) of events occurring in the virtual space. Since the avatar has an abstract shape, it can use only abstract methods (based on emotional associative techniques) to express its state. This test evaluates how well the avatar's expression is understood.*
> *In the following sections, we will present you with 4 tasks, each consisting of 4 videos. The first 2 videos in each task will show user interactions in the virtual space, while the third and fourth video will show the avatar's response to user behaviors from one of the user videos. Your goal will be to pair the first two (user) videos with the second two (avatar) videos.*
> *We used MaxWhere VR application to create the virtual environment for this experiment. This application allows users to access or upload media contents through interactive boards in the 3D space, also known as "smartboards". The user interaction patterns in this system can be classified into three types:*
> - navigation: moving in the space
> - 2D interaction: viewing content on smartboards
> - manipulation: moving, resizing smartboards
>
> *We recorded user interaction videos in this study that include these patterns. These patterns affect the "mood" of the abstract avatar, which changes the displayed overall state of the environment to combination of these values: "bored", "excited", "happy" or "frustrated". In the following tasks, the overall state will be the* "usefulness" *of user interactions, which means how much the user takes advantage of every function available in the space. For instance,*
> - a 'pointless wandering' pattern will trigger a frustrated mood,
> - a task pattern will trigger a combination of bored mood and excitement based on how repetitive is the task,
> - and a balanced use of all interaction types will trigger a happy mood.
>
> *The recordings are about 30 seconds long and each are accelerated to the same degree.*"

We also present a brief video to the participants after the introduction, which demonstrates the typical user interactions in MaxWhere VR that are described in the briefing text. This is essential for ensuring that they can recognize and differentiate the actions on user interaction videos with as much confidence as a more experienced MaxWhere user.

C. *Automated data collection*

The survey used an automatic recording system to store each response in a spreadsheet that was linked to the survey. The recording also triggered several custom routines that analyzed the answer and produced some statistical metrics. To do this, we created a custom scripting interface that connected the survey with the data collector modules that were integrated into the survey. The custom routines generated the survey for the next participant. This way, the next version of the test had different order and combinations of answers and videos.

The survey for the experiment is created in Google Forms linked to a Google spreadsheet to store the results. Custom routines were implemented in Google Apps Script framework to process the data. The script runs automatically after each Form submission, triggered by the framework settings.

D. *Custom routines*

Custom routines performed upon survey completion includes randomization of the next test and evaluation of the current one.

Video randomization is performed using the following steps in the script.

- Read out the URLs of video pairs from a spreadsheet page. Each video pair (user interaction and avatar video) is associated with an ID.
- Associate groups to the video pairs according to which user interaction type they contain
- For the first task, select the Idle video and pair it with one of the one-dimensional videos (Discovery, Manipulation, Operation2D).

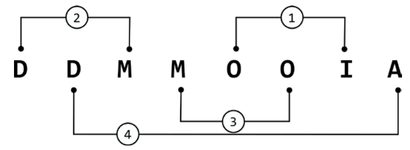


Fig. 4. Example run of the randomization routine. Numbers depicts the order of operation. Letters are the user interaction types (D - discovery, M - manipulation, O - 2D operation, I - idle, A - all)

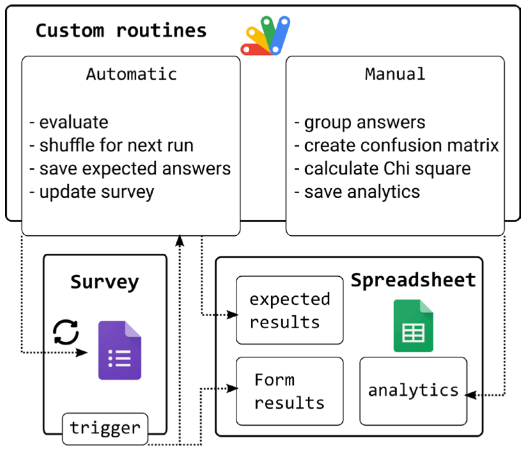


Fig. 5. Data flow of survey, connected spreadsheet and custom routines. Automatic routines run when a survey is completed. It evaluates current answers and shuffle pairs for a new run. Manual routines run on demand and create analytics.

- For second and third task, pair the least used one-dimensional videos randomly. Least used videos are determined by counting how many times a video is already used for pairing in previous tasks. (thus, it is actually forced for the second task but not for the third)
- Pair the All dimensional video with the remaining video for the last (fourth) task.
- Randomize order within pairings to alter display order of videos in the survey.

The outcome of random pairing for the tasks of the next survey is then registered in a spreadsheet. It represents the expected answers for the next submission. An example run of the randomization routine is demonstrated on Figure 4.

The form submission trigger runs the script after the connected spreadsheet has been updated with the latest responses. This allows the script to automatically check the accuracy of the matching tasks. The script retrieves the correct answers from the sheet based on the latest response.

The overall functionality of the evaluation and randomization script can be seen on Figure 5.

VI. Results

In this section we summarize the statistical analysis of the survey responses. The main objective of the analysis is to determine how well the users can match the avatar behaviors with the corresponding user interactions.

The survey has been distributed only for a smaller group at first for fine-tuning. Based on the preliminary findings described below, we modified the survey and recruited a larger group of participants through social media and technical forums. We analyzed the final responses using the custom manual methods explained in the previous section.

A. Preliminary results

For the test video pairs, we initially had a slightly different setup. In the first version of the experiment we combined 2 different interaction types as well instead of using only one, zero or all dimensions for creating test videos. The randomization algorithm also did less instructed shuffling, which made the video match tasks more challenging.

We recorded answers from 12 people (5 male, 7 female) and found that none of the interaction types produced acceptable recognisability. We also observed that the randomized task generator created complex combinations that increased the difficulty level for even the most skilled MaxWhere users.

B. Revised experiments

Based on the preliminary results we simplified the experiment tasks by reducing the number of user interaction types depicted in the test videos. This made the videos more distinguishable as the participants did not see tasks with combinations of the same user interaction type.

We also modified the shuffle algorithm in order to fix one half of the random pairings in half of the tasks. The first task has always presented the zero dimensional case and the last task had all the user interaction types for one of the video pairs. The revised video pairing is described in steps in section V-D. The new routine solved the distribution problems as well, because each video pair is used once for each participant.

The explanatory text at the beginning of the survey has also been modified in order to provide a better understanding of the upcoming tasks. We added information about MaxWhere VR in general and introduced the typical user interaction types performed in the user videos of the tasks in a bullet point list and on a short video. We emphasised that the aggregated state of the space is the "usefulness" of user interactions in this experiment and listed several examples for mappings between the user interaction and displayed emotions of the avatar. For example, we described that a happy avatar means that the user interaction was useful, while a frustrated or bored avatar means that it was not. These sentences helped the participants to understand the situation presented on user interaction videos and emotional states shown on avatar videos. We considered this to be reasonable for our experiment, realizing that it would have been difficult to expect users to associate interactions with the avatar if they had not been attuned to the kinds of interactions that exist in the first place, or the general purpose of the avatar.

C. Statistical analysis

We recorded responses from 31 participants (8 female, 22 male, 1 undefined). Most of them were middle-aged with average age between 30 and 41. Based on the media channels we distributed the survey, we could ascertain that participants

TESTS	Discover	Manipulate	Operate2D	Idle	All	MATCHES	Discover	Manipulate	Operate2D	Idle	All
Discover	0	20	23	8	11	Discover	0	14	12	7	7
Manipulate	20	0	19	14	9	Manipulate	14	0	12	11	7
Operate2D	23	19	0	9	11	Operate2D	12	12	0	7	7
Idle	8	14	9	0	0	Idle	7	11	7	0	0
All	11	9	11	0	0	All	7	7	7	0	0
EXPECTED	Discover	Manipulate	Operate2D	Idle	All	OBSERVED	Discover	Manipulate	Operate2D	Idle	All
Discover	31	10	11.5	4	5.5	Discover	40	6	11	1	4
Manipulate	10	31	9.5	7	4.5	Manipulate	6	44	7	3	2
Operate2D	11.5	9.5	31	4.5	5.5	Operate2D	11	7	38	2	4
Idle	4	7	4.5	15.5	1.00E-10	Idle	1	3	2	25	0
All	5.5	4.5	5.5	1.00E-10	15.5	All	4	2	4	0	21

Fig. 6. Tables used for extracting input data for Chi Square test. Tests table consists of participation of interaction types in tests. Matches table shows the successful matches of the user interaction types. Expected and Observed tables are calculated for Chi Square test using the basic assumption that user interaction types can be recognized only as good as 50%

Confusion / Occurences	Discover	Manipulate	Operate2D	Idle	All
Discover		0.7	0.522	0.875	0.636
Manipulate	20		0.632	0.786	0.778
Operate2D	23	19		0.778	0.636
Idle	8	14	9		0
All	11	9	11	0	

Fig. 7. Confusion and occurrence matrix of the matching results. The matrix has two parts: the lower left part shows the total number of appearances of each user interaction type combination, and the upper right part shows the rate of successful matches for each combination. Darkness of background color of cells indicates the significance of values

CHIDIST	Discover	Manipulate	Operate2D	Idle	All
p values	0.141611638	0.02256998518	0.3981787884	0.0193337095	0.3849561179

Fig. 8. P values of Chi Square test. The result shows significant difference for Idle and Manipulate state and unconfirmed difference for Discover and Operate2D.

had adequate computer skills and at least half of them were familiar with 3D video games.

We performed various statistical tests and analyses to assess how well our avatar can convey the aggregated state of the space during different types of user interaction. To obtain meaningful statistics, we first grouped the data by dimensions of user interaction type. We counted the successful matches from the tasks for every participant and registered a successful recognition of the aggregated state for each successful match. Since the matching tasks involved two different dimensions each time, we recorded success or failure for both of the involved interaction types.

We constructed two basic tables from the extracted data: The total number of participations and number of successful matches for each user interaction dimension. Both tables were derived from the spreadsheet that contained the pairing information about each test run. Using the pairing information alone we could construct the participation matrix. The automated evaluation script registered the success for each task of each test run, which was used to construct the correct matches matrix. Results of extracted tables can be seen on Figure 6.

Based on the participation and correct match tables we constructed a confusion matrix. This matrix shows how well the participants matched the items correctly in each task. It helps us to evaluate how well the expressive ability of our avatar performs for different types of user interactions. We obtained the accuracy by dividing the number of correct matches by the number of total appearances. Figure 7 shows the confusion matrix we generated. The Idle and Discover scenario had the highest type combination score. This makes sense because the avatar shows anger in one case and boredom in the other.

For our statistical test we used Chi square test to see how the aggregated avatar performed. A Chi square test is a statistical method that can be used to test the association between pairs of categorical variables. The test compares the observed frequencies of each category with the expected frequencies under the null hypothesis of no association. The p-value of the test is the probability of obtaining a test statistic as extreme or more extreme than the observed one, under the null hypothesis. A small p-value indicates that there is strong evidence to reject the null hypothesis and conclude that there is an association between the variables – in our case, that users peformed better than randomly in the case of at least one category.

As a baseline, or null hypothesis we used the assumption that the avatar does not help in recognition of the aggregated state of the virtual environment. This means that we suggested that the participants choose randomly during the matching tasks giving us 50% success rate for each user interaction types.

Therefore, we set the expected table for Chi square test reflecting this statement. Each value in the matrix has been calculated by taking the half of the total occurrences of the given type combination. We used the matrix diagonal to store the summed result for a given dimension. For example, the total expected value for Discover user interaction type is the sum of each test combination containing Discovery divided by two, which is 62 / 2. Note, that we also simplified the case of zero values in the expected matrix to keep the calculation straightforward. We replaced zeros (which would cause a divide error during the upcoming calculations) with 1E-10 values. The expected table can be seen in lower left side of Figure 6.

We obtained the observed values for the Chi square test from the correct match matrix. In the diagonal we use the number of observed successful matches for a given interaction type which is calculated by the sum of correct matches for a given dimension in each combination it participates. For non-diagonal cells we used the number of *incorrect* choices of a given interaction type combination as the expected values describes the choice of the given value *despite of* the correct answer. We calculated these values by subtracting the correct match value for each combination from the total number of occurrences. For example, see the first two cell of the Observed matrix on Figure 6. The observed total of successful matches of Discover dimension is 14 + 12 + 7 + 7 = 40. The total number of times when participants mistakenly chose Manipulate instead of Discover when these two dimensions were combined is 20 - 14 = 6

Using the composed matrices we could calculate the P values of the Chi square test for each user interaction type as shown on Figure 8. We used 5% for the statistical threshold to reject the null hypothesis.

Based on the results we can state that Idle and Manipulate state show extreme deviation from the values picked by chance (Manipulate 2.2%, Idle 1.9%), thus the effect of our avatar for choosing the correct matching were statistically significant. With other words it is extremely improbable that the results are unrelated to the avatar setup. All and Operate types showed no real deviation from the randomly selected answers thus they could not produce evidence for the usefulness of the avatar. Discover type however shows high probability for existing effect of the avatar without being considered as significant result (14% chance for no connection).

VII. Conclusions

In this paper, we broadened the scope of the avatar concept to include multi-user avatars using an abstract visual language, referred to as aggregated avatars. We argued that such avatars could be useful in multi-user platforms such as VR-based working environments. We proposed a design methodology and developed a reference implementation of the aggregated avatar concept on the MaxWhere VR platform.

Based on our statistical analyses we can conclude that two major types of user interactions (idle state, and rearrangement of spatial content) could be very well recognized by users with the use of our abstract aggregated avatar. A third type of interaction, referring to users moving around in the space, were also likely to be recognized by users at a higher success rate than chance. Other dimensions provided a promising but statistically not relevant result – although in the case of the “All” state, this could have been due to the fact that this state included some examples of all other interactions. Based on these results, we conclude that the proposed avatar design could be applicable to some contexts, and is worthy of further study and refinement.

VIII. Acknowledgement

Project no. C1015653 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the KDP-2020 funding scheme.

This research is contributing to project no. 2021-1.1.4-GYORSÍTÓSÁV-2022-00081 that has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021-1.1.4-GYORSÍTÓSÁV funding scheme.

The research presented in this paper was also supported by the Hungarian Research Network (HUN-REN), and was partly carried out within the HUN-REN Cognitive Mapping of Decision Support Systems research group.

References

[1] L. de Wildt, T. H. Apperley, J. Clemens, R. Fordyce, and S. Mukherjee, “(Re-) Orienting the Video Game Avatar,” *Games and Culture*, vol. 15, no. 8, pp. 962–981, 2020. DOI: 10.1177/1555412019858890.

[2] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch, “The Effect of Avatar Realism in Immersive Social Virtual Realities,” in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10, 2017. DOI: 10.1145/3139131.3139156.

[3] F. Sibilla and T. Mancini, “I Am (Not) My Avatar: A Review of the User-Avatar Relationships in Massively Multiplayer Online Worlds,” *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, vol. 12, no. 3, 2018. DOI: 10.5817/CP2018-3-4.

[4] P. Baranyi, A. Csapo, and G. Sallai, *Cognitive Infocommunications (CogInfoCom)*. Springer, 2015. DOI: 10.1007/978-3-319-19608-4.

[5] I. Horváth, Á. B. Csapó, B. Berki, A. Sudár, and P. Baranyi, “Definition, Background and Research Perspectives Behind ‘Cognitive Aspects of Virtual Reality’(cVR),” *Infocommunications Journal*, no. SP, pp. 9–14, 2023. DOI: 10.36244/ICJ.2023.SI-IODCR.2.

[6] D. Aneja, D. McDuff, and S. Shah, “A High-Fidelity Open Embodied Avatar with Lip Syncing and Expression Capabilities,” in *2019 International Conference on Multimodal Interaction*, pp. 69–73, 2019. DOI: 10.1145/3340555.3353744.

[7] D. Roth, J.-L. Lugrin, D. Galakhov, A. Hofmann, G. Bente, M. E. Latoschik, and A. Fuhrmann, “Avatar Realism and Social Interaction Quality in Virtual Reality,” in *2016 IEEE Virtual Reality (VR)*, pp. 277– 278, IEEE, 2016. DOI: 10.1109/VR.2016.7504761.

[8] T. Alldieck, M. Magnor, W. Xu, C. Theobalt, and G. Pons-Moll, “Detailed Human Avatars from Monocular Video,” in *2018 International Conference on 3D Vision (3DV)*, pp. 98–109, IEEE, 2018. DOI: 10.1109/3DV.2018.00022.

[9] L. Hu, S. Saito, L. Wei, K. Nagano, J. Seo, J. Fursund, I. Sadeghi, C. Sun, Y.-C. Chen, and H. Li, “Avatar Digitization from a Single Image for Real-Time Rendering,” *ACM Transactions on Graphics (ToG)*, vol. 36, no. 6, pp. 1–14, 2017. DOI: 10.1145/3130800.31310887.

[10] J. L. Ponton, V. Ceballos Inza, L. Acosta, A. Rios, E. Monclús, and N. Pelechano, "Fitted Avatars: Automatic Skeleton Adjustment for Self- Avatars in Virtual Reality," *Virtual Reality*, vol. 27, pp. 1–20, 07 2023. DOI: 10.1007/s10055-023-00821-z.

[11] C. Pedica and H. Högni Vilhjálmsson, "Spontaneous Avatar Behavior for Human Territoriality," *Applied Artificial Intelligence*, vol. 24, no. 6, pp. 575–593, 2010.DOI: 10.1080/08839514.2010.492165.

[12] A. Kleinsmith, P. R. De Silva, and N. Bianchi-Berthouze, "Cross-Cultural Differences in Recognizing Affect from Body Posture," *Interacting with computers*, vol. 18, no. 6, pp. 1371–1389, 2006. DOI: 10.1016/j.intcom.2006.04.003.

[13] M. Fabri, D. J. Moore, and D. J. Hobbs, "The Emotional Avatar: Non-Verbal Communication between Inhabitants of Collaborative Virtual Environments," in *Gesture-Based Communication in Human-Computer Interaction: International GestureWorkshop, GW'99 Gif-sur-Yvette, France, March 17-19, 1999 Proceedings*, pp. 269–273, Springer, 1999. DOI: 10.1007/3-540-46616-9_24.

[14] J. Yang, R. T. Marler, H. Kim, J. Arora, and K. Abdel-Malek, "Multi-Objective Optimization for Upper Body Posture Prediction," in *10th AIAA/ISSMO multidisciplinary analysis and optimization conference*, p. 4506, 2004. DOI: 10.2514/6.2004-4506.

[15] R. Rivu, D. Roth, F. Alt, and Y. Abdelrahman, "The Influence of Avatar Personalization on Emotions in VR," *Multimodal Technologies and Interaction*, vol. 7, p. 38, 03 2023. DOI: 10.3390/mti7040038.

[16] G. Persa and Á. B. Csapó, "A Fuzzy Driven Framework for the Cognitive Modeling and Quantitative Analysis of Aggregated Avatars," in *CINTI-MACRo 2022*, pp. 000 143–000 150, IEEE, 2022. DOI: 10.1109/CINTI-MACRo57952.2022.10029574.

[17] P. S. Moyer, G. Salkind, and J. J. Bolyard, "Virtual Manipulatives used by K-8 Teachers for Mathematics Instruction: The Influence of Mathematical, Cognitive, and Pedagogical Fidelity," *Contemporary Issues in Technology and Teacher Education*, vol. 8, no. 3, pp. 202–218, 2008.

[18] B. Bos, "Virtual Math Objects with Pedagogical, Mathematical, and Cognitive Fidelity," *Computers in Human Behavior*, vol. 25, no. 2, pp. 521–528, 2009. DOI: 10.1016/j.chb.2008.11.002.

[19] D. Lansel, T. Kergoat, S. Bardin, and L. Deruy, "Using Virtual Reality Domain to Set Up Sports Simulation as Part of Rehabilitation," 2021.

[20] Á. B. Csapó, I. Horvath, P. Galambos, and P. Baranyi, "VRasa Medium of Communication: From Memory Palaces to Comprehensive Memory Management," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 000 389–000 394, IEEE, 2018. DOI: 10.1109/CogInfoCom.2018.8639896.

[21] G. Persa, A. Csapo, and P. Baranyi, "CogInfoCom Systems from an Interaction Perspective – A Pilot Application for EtoCom," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 16, no. 2, pp. 297–304, 2012. DOI: 10.20965/jaciii.2012.p0297.

[22] B. Korcsok, V. Konok, G. Persa, T. Faragó, M. Niitsuma, Á. Miklósi, P. Korondi, P. Baranyi, and M. Gácsi, "Biologically Inspired Emotional Expressions for Artificial Agents," *Frontiers in psychology*, vol. 9, p. 1191, 2018. DOI: 10.3389/fpsyg.2018.01191.

[23] A. Sudár and Á. B. Csapó, "Descriptive Markers for the Cognitive Profiling of Desktop 3D Spaces," *Electronics*, vol. 12, no. 2, p. 448, 2023. DOI: 10.3390/electronics12020448.

[24] B. Berki, "Does Effective Use of MaxWhere VR Relate to the Individual Spatial Memory and Mental Rotation Skills?," *Acta Polytechnica Hungarica*, vol. 16, no. 6, pp. 41–53, 2019. DOI: 10.12700/APH.16.6.2019.6.4.

[25] G. Stankov, B. Nagy, *et al.*, "Eye Tracking Based Usability Evaluation of the MaxWhere Virtual Space in a Search Task," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 469–474, IEEE, 2019. DOI: 10.1109/CogInfoCom47531.2019.9089922.

György Persa This author obtained his MSc degree in 2009 from Pázmány Péter Catholic University in Budapest.
Following his graduation, he conducted scientific research at the Institute for Computer Science and Control of the Hungarian Academy of Sciences.
In 2020, he began his doctoral studies at Széchenyi István University in Győr. His primary research focus is on design and implementation in the areas of virtual reality, 3D applications, avatars, human-machine interfaces and cognitive computing.

Ádám B. Csapó obtained his PhD degree at the Budapest University of Technology and Economics in 2014. From 2016 to 2022, he has worked as an associate professor at the Széchenyi István University in Győr; and, since 2022, he continued his academic journey as as an associate professor at Óbuda University, Budapest, Hungary. Since September 2023, he has held the position of associate professor at Corvinus University of Budapest. Dr. Csapó's research focuses on cognitive infocommunication channels in virtual collaboration environments, i.e. enabling users to communicate with each other and with their spatial surroundings in novel and effective ways. At the same time, he has been involved in the development of assistive technology for the visually impaired, as well as in the development of a commercial VR platform. Dr. Csapó has over 50 publications, including 1 co-authored book and 20 journal papers, and has actively participated in the organization of numerous international conferences and special issues.

Special Issue
of the Infocommunication Journal

Cognitive Aspect of Emotion Estimation of a Driver

Jinshan Luo, Yuki Okaniwa, Natsuno Yamazaki, Yuko Hiramatsu, Madoka Hasegawa, *Member, IEEE* and Atsushi Ito *Member, IEEE*

***Abstract*—Despite rapid advancements in the automotive industry, traffic safety risks persist. Addressing this challenge requires innovative driver assistance technologies. Common accidents result from driver inattention, fatigue, and stress, leading to issues like falling asleep at the wheel and improper acceleration and braking. Our study aims to contribute to advanced driver assistance systems that adapt to drivers' emotional needs, ultimately enhancing road safety. In this paper, we mention the result of our research to estimate drivers' emotions using sensors. For that purpose, we developed a sensor network containing sensors such as EEG, eye tracker, and driving simulator. We explored the relationship. As a result, we confirmed the relation between the driver's emotions, especially sleep conditions, driving speed, duration, and brain wave behavior.**

***Index Terms*—driving support system, emotion estimation, EEG, eye tracking, sensor network, persistent homology**

I. Introduction

Amidst the rapid devolution of the automotive industry, the persistence of traffic safety risks remains a critical concern. Consequently, the development of intelligent and innovative driver assistance technologies becomes imperative. Among the factors contributing to traffic accidents, the most prevalent include falling asleep at the wheel, improper acceleration, and braking errors. These mishaps stem from driver inattention, fatigue, drowsiness, and stress, all significantly contributing to road accidents. In particular, the cumulative effects of fatigue resulting from extended periods of driving can substantially impair a driver's cognitive function and judgment, ultimately leading to drowsiness. Hence, it becomes essential to promptly detect driver fatigue and drowsiness during driving and institute appropriate measures. To address this challenge, there is a pressing need to develop support systems that encourage drivers to take breaks and enhance the in-vehicle environment to optimize comfort.

Given this backdrop, our research endeavors to objectively assess driver fatigue by gathering an extensive array of data, encompassing ElectroEncephaloGraphy (EEG), heart rate, eye movements, and driving activities, all obtained through driving simulators. We aim to scrutinize this data meticulously, with the expectation of delivering valuable insights that can enhance driver safety and mitigate the risk of traffic accidents.

It is worth noting, however, that while EEG holds promise as an excellent indicator of emotional states, practical limitations arise when attempting to employ EEG sensors during driving. Additionally, EEG sensors may provide limited accuracy in measuring neural activity related to emotions occurring beyond the upper layers of the brain. [1].

The remaining part of this paper is structured as follows. Section 2 introduces a variety of related research. We explain the research goal, problems, and objectives in Section 3. The sensors and network we used in the research are explained in Section 4. Section 5 explains the detailed experiment result, and discuss and analyze the result in Section 6. Finally, we present our conclusions in Section 7.

II. Positioning and the Related Works of the Paper

Numerous studies have been undertaken to predict driver fatigue by examining the correlation between drivers' biological signals and their eye movements.

A. Relationship between this study and Cognitive Infocommunications

This paper proposes and uses a sensor system to estimate a driver's mind state and subsequently explore the relationship between emotional states and driving behavior. This study combines artificial and natural cognitive capabilities. The whole system's new hybrid cognitive capabilities fall into the concept of Cognitive Infocommunications [2,3].

One of the branches of Cognitive Infocommunications focuses on Cognitive Mobility, which investigates the entangled combination of research areas such as mobility, transportation, vehicle engineering, social sciences, artificial intelligence, and Cognitive Infocommunications [4,5].

Thus the overall new capability of the combination of the censoring system and the driver leads to a new capability of the whole system that is to improve the driving effectiveness to avoid accidents and further car design outcomes.

This work was supported in part by JSPS Kakenhi under Grants, JP17H02249, 18K111849, 20H01278, 20H05702, 22K12598, 23H03649.

Jinshan Luo is with Toyota Customizing and Development Co., Ltd, Ashigara-gun, Kanagawa, Japan (e-mail: dc227209@s.utsunomiya-u.ac.jp).

Yuki Okaniwa is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail: a21.hcsk@g.chuo-u.ac.jp).

Natsuki Yamazaki is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail: a22.x3yf@g.chuo-u.ac.jp).

Yuko Hiramatsu is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail: susana_y@tamacc.chuo-u.ac.jp).

Madoka Hasegawa is with the School of Engineering Utsunomiya University, Tochigi, Japan (e-mail: madoka@is.utsunomiya-u.ac.jp).

Atsushi Ito is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail:atc.00s@g.chuo-u.ac.jp).

DOI: 10.36244/ICJ.2024.5.12

B. Studies Using Heart Rate and Electrocardiogram [6]

This study reported that during the transition from mental fatigue to drowsiness, the number of blinks increased and, conversely, the heart rate decreased.

C. Studies utilizing gaze angle and eye rotation angle [7]

This study used image sensors to generate computational models from gaze and eye rotation angles. As a result, it was reported that the accuracy of measuring mental workload from driving was improved from eye movements.

D. Studies Using Gaze Angle and Eye Rotation Angle [8].

This study has shown that blinking decreased with task load during driving. However, another paper [9] has found the opposite result, that blinking increases with increasing task load, and there still needs to be a clear answer regarding how it can be used.

Furthermore, numerous studies have explored the monitoring of driver emotions using sensors as a means to prevent risk-taking behavior. These investigations involved using multiple driving simulators to simulate realistic driver interactions and estimated driver emotions based on driving performance data. However, these studies did not directly detect emotional data to validate the accuracy of their measurements [10]. In [11], researchers developed a sensor network to establish a mapping relationship among various sensor data to monitor driver emotions. This approach incorporated metrics such as heart rate, skin conductance, skin temperature, and facial expressions. Nevertheless, the research did not address driving performance data closely tied to driving behavior. [12] introduced a non-intrusive emotion recognition system designed for car drivers, employing a thermal camera to enhance Advanced Driver Assistance Systems (ADAS). However, it's important to note that this system has yet to be tested in actual driving conditions, which leaves room for further exploration and validation.

Prior research efforts have explored various avenues in the realm of emotion recognition for drivers. For instance, in [13], an approach centered around facial expressions was introduced. This approach leveraged a comprehensive on-road driver facial expression dataset, encompassing diverse road scenarios and corresponding driver facial expressions during driving. Meanwhile, [14] devised a methodology that combines Local Binary Pattern (LBP) features with facial landmark features to detect driver emotions. This method further employed a supervised machine learning algorithm, specifically a support vector machine, to classify different emotions effectively.

Additionally, [15] put forward an innovative approach, introducing a custom-created Convolutional Neural Network (CNN) feature learning block to enhance the performance of an existing 11-layer CNN model. This augmentation resulted in an improved and faster R-CNN face detector capable of accurately identifying the driver's face. However, it's essential to note that these studies primarily focused on processing facial image data for driver emotion recognition. They did not delve into aspects such as body motion or explore the intricate relationship between driver emotion and driving performance data.

III. Outline of this study

A. Research Goal

This research aims to estimate drivers' emotions during driving to prevent car accidents.

B. Problems and Objectives of the research

While several studies have explored the estimation of driver emotions through means such as brain waves and other bio-signals, there are two notable challenges to consider. Firstly, relying solely on EEG may be problematic due to potential variations caused by the experimental environment, introducing an element of risk. Secondly, the practicality of measuring EEG by having drivers wear sensors while driving is a concern.

To address these issues, we aim to identify alternative sensors that are easy to use, robust, and cost-effective. We will compare these potential sensors with popular vital sensors commonly employed for health monitoring, as well as sensors integrated into vehicles but not worn by individuals. By examining these options, we can explore the feasibility of replacing EEG with more practical sensor solutions for emotion detection in a driving context.

IV. Sensors

We implemented a sensor network for driver emotion monitoring around a driving simulator. In this section, we would like to explain each sensor and outline the network.

A. Driving simulator

To achieve the goal of estimating the driver's emotions by using a drive recorder and analyzing the relation between driver emotions and behavior during driving, we collect driving performance data from the driving simulator such as speed, accelerator pedal degree, brake pedal pressure, steering angle, and distance from the start.

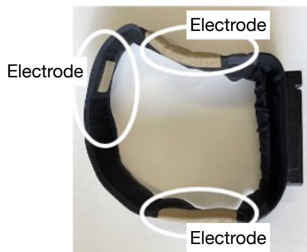


Fig. 1. EEG sensor (headband) [13]

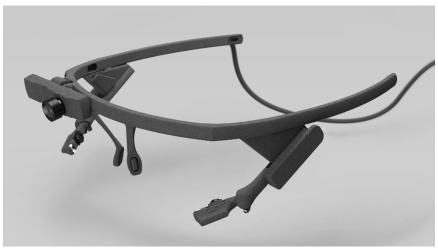


Fig. 2 Eye tracking device [16]

B. EEG sensor

We used an EEG sensor to estimate the mind state of a driver. Typically, EEG sensors are large devices with many electrodes used in hospitals, but using such devices in a car or a driving simulator is not easy [16]. So, we have developed a wearable EEG sensor that can acquire data for several hours without stress, as shown in Figure 1. This EEG sensor has BLE and sends data in real time. The sampling rate is 512 Hz. In this experiment, we used M5Stack Core2 [18] as a receiver of the EEG signal, and the data is written on an SD card in the receiver device. The EEG sensor can output the information listed in Table 1. In addition, our sensor can output two additional information: Attention (concentration, similar to Beta wave) and Mediation (relaxing, similar to Alpha wave) [19].

C. Eye tracker

We utilized an eye tracker, specifically the Pupil Core [20], to monitor eye movements accurately, as illustrated in Figure 2. This device provides us with the precise x and y coordinates of the gaze, enabling us to simultaneously capture video footage of the surrounding scenery and the movement of the eyeball.

D. Network

As shown in Figure 3, we established a sensor network for data collection, with certain components connected via Bluetooth Low Energy (BLE [21]) for real-time data transmission. In contrast, other components, such as sensors connected on Controller Area Network (CAN [22]) in the driving simulator, remained offline for security considerations.

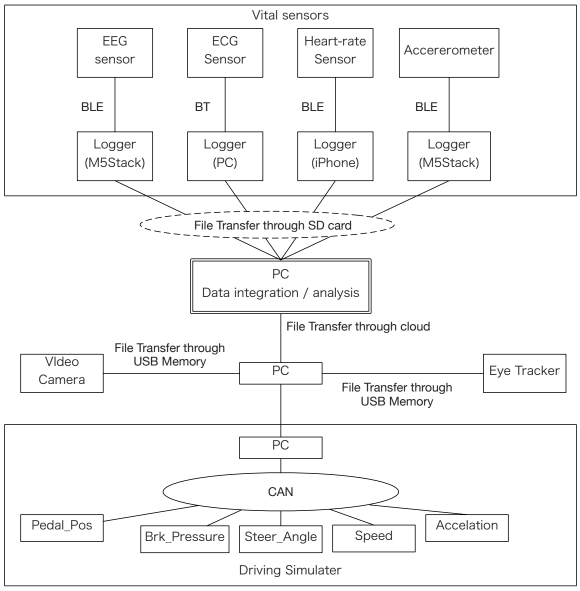


Fig. 3. The system diagram

TABLE II
BRAIN WAVES [19].

Frequency band name	Frequency	Brain states
Delta	0.5–4 Hz	Sleep
Theta	4–8 Hz	Deeply relaxed, inward focused
Alpha	8–12 Hz	Very relaxed, passive attention
Beta	12–35 Hz	Anxiety dominant, active, external attention, relaxed
Gamma	Over35 Hz	Concentration

V. EXPERIMENT AND RESULT ANALYSIS

A. Experiment Design

We set two test courses in the driving simulator. One is Tokyo Metropolitan Highway (C1), and another is a road in the center of Paris. The details of the setting are shown in Table 1. In the C1 course, we changed the brightness during driving from daytime to evening. In Paris, we used Simulation of Urban Mobility (SUMO) [23] to provide some interference to drivers, such as traffic and unexpected behavior of pedestrians.

On August 9 and 21, 2023, and November 22, 2023, two students of Chuo University (20 years old, and 2years old, owning a driver's license) drove C1 and Paris. The duration of each driving test was 45 minutes. As per our experiences, after 30 minutes, a driver starts to feel fatigued, so we set 45 minutes. In this paper, we label each trial as 20230809-C1-1, C1-2, Paris-1, Paris-2, 20230821-C1-1, C1-2, Paris-1, Paris-2, and 20231122-Paris-5.

Before and after driving, the test driver answered PANAS (The Positive and Negative Affect Schedule [24]) with 10 Positive and 10 Negative questions to record the mental state. As shown in Figure 4, the score of negative questions increased in all cases, indicating that the test driver consistently reported feeling fatigued after 45 minutes of driving.

B. Fatigue from distance and time

Figure 5 presents the relationship between the number of rounds and the duration of a single round of driving. Notably, it becomes evident that, after several rounds, the lap time increased by approximately 20%. This observation suggests

TABLE I
TEST COURSE IN THE DRIVING SIMULATOR.

Course	Time/round	SUMO	Brightness change (in 45 min)
Tokyo Metropolitan Highway C1	About 10 min	No	0-10 min 4:00 PM, 10-20 min 6:00 PM 20-30 min 7:00 PM, 30-40 min 7:30 PM 40-45 min 8:00 PM (with road lighting)
Paris City Area Course	About 6 min	Traffic and Pedestrian crossing road	No

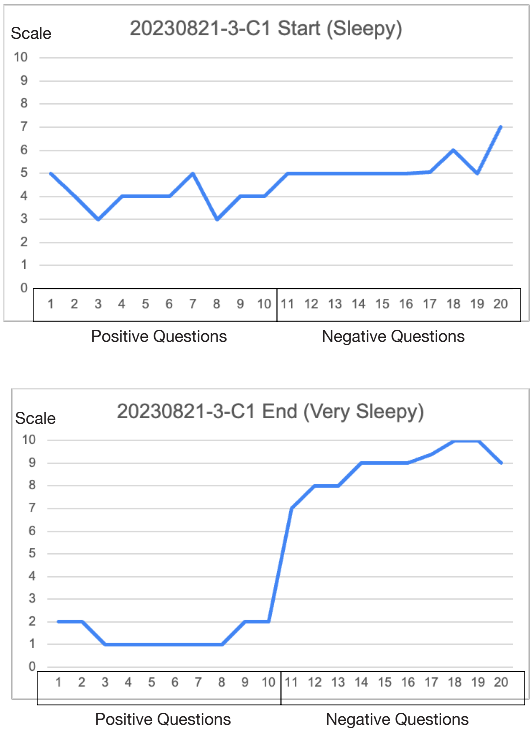


Fig. 4. Example of the result of PANAS

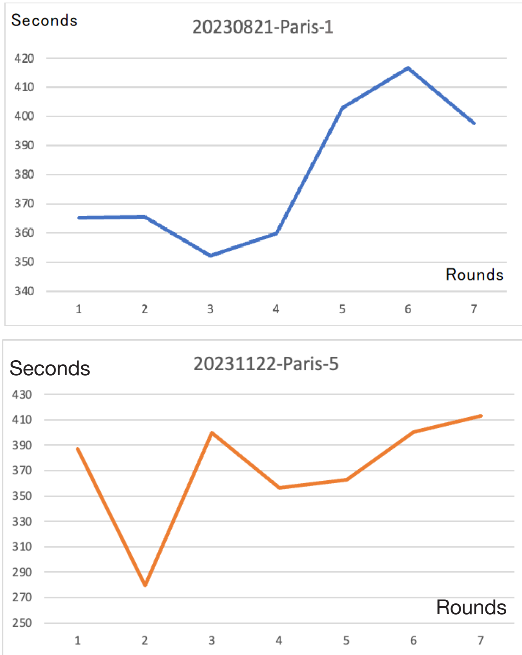


Fig. 5. Example of rap time of each round

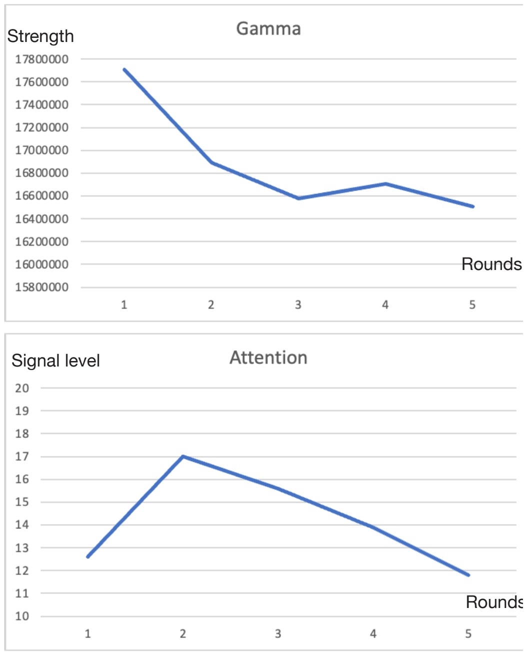


Fig.6. Trend of Gamma and Attention of 20230821-C1-1 (going down along the timeline)

that a driver's concentration tends to decline after experiencing fatigue, resulting in a slowdown in driving speed.

C. Fatigue from changing brightness

Figure 6 illustrates the connection between brightness levels and changes in emotion detected through EEG. As the driving rounds progress, the road becomes darker, and a corresponding decrease in driver concentration is evident. This pattern is consistent with the trends observed in C1 driving data. Conversely, during driving experiences in Paris, there was no significant decline in attention levels. Hence, we can infer that darkness has a detrimental impact on a driver's attentiveness, potentially contributing to increased fatigue and decreased concentration during nighttime driving scenarios.

D. Relation between EEG and eye tracking

Figure 7 illustrates the link between Gamma brainwaves and eye blinks. Our analysis of Paris data revealed a consistent pattern: when Gamma fell below 1.0E7, indicating reduced brainwave activity, the driver often lost concentration, leading to eye blinks or closures (below the red line in Figure 7). This suggests Gamma changes are a valuable indicator of tiredness, especially sleepiness, aligning with the driver's drowsiness in the latter part of the round. Additionally, this finding underscores the significance of eye tracking as a reliable method for monitoring the driver's movements and quantifying their level of fatigue while actively engaged in driving.

E. Relation between facial recognition and eye tracking

Facial expressions directly reflect emotions, and body motion strongly associates with emotions [25]. To estimate a driver's stress and fatigue, we harnessed the effectiveness of a widely used drive recorder, capable of capturing facial expressions and body movements. Consequently, in this study, we amalgamated driving performance data from a driving simulator with facial expression and body motion data obtained from a drive recorder for a comprehensive correlation analysis.

For facial expression recognition and head motion measurements, we utilized MediaPipe Face Mesh [26], with

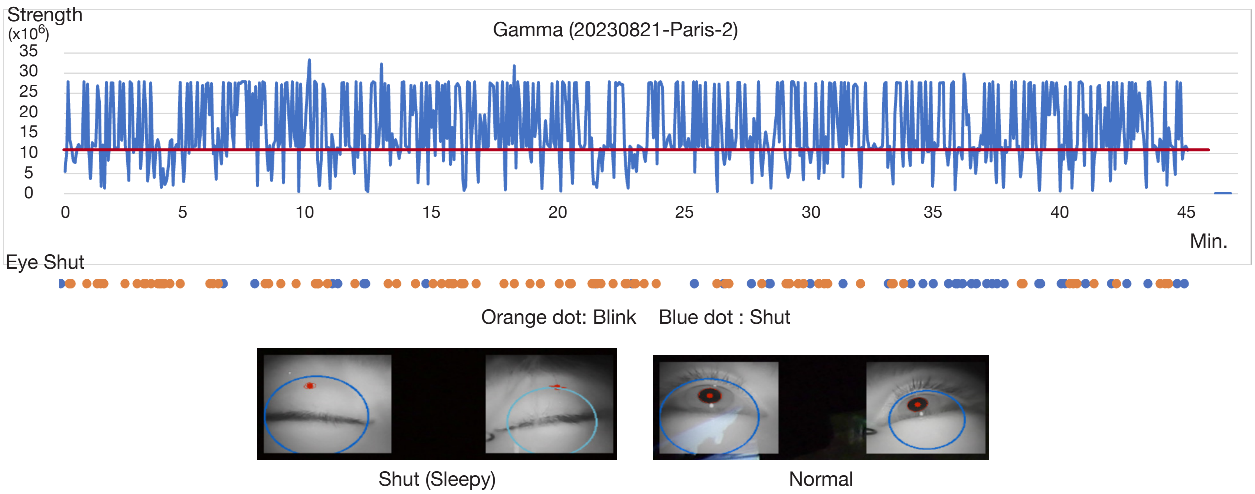


Fig. 7 Relation between Gamma wave and eye blink

prior video image preprocessing carried out using OpenCV [27]. This allowed us to visualize driving performance data concerning the accelerator pedal, brake, and steering, shedding light on their relationship with concurrent facial expressions and body motions. To assess the drive recorder's performance, we conducted a comparative analysis of emotion measurement results obtained through the drive recorder and those from an EEG sensor as our previous study [28]. This approach facilitates an evaluation of the drive recorder's effectiveness in gauging driver emotions and provides a low-cost and easily implementable method for collecting data on drivers' facial expressions through video footage was proposed.

We analyzed the connection between facial recognition and eye tracking, with a specific emphasis on the occurrence of eye closures. In Figure 8, we present six distinct patterns of facial classification observed during the experiment, which, in turn, allow us to infer four primary emotions: (1) Neutral, (2) Anxiety, (3) Boredom, and (4) Fatigue. Our primary focus lies on fatigue as it relates to the sensation of tiredness while driving. Table 3 presents the number of reported fatigue feelings and occurrences of eye closures per 5-minute intervals. The low p-value obtained from the T-Test (0.4) further validates the strong relationship between eye movements and facial expressions, supporting our hypothesis that eye tracking effectively correlates with driver emotion, particularly in instances of fatigue.

VI. Discussion

This paper aims to estimate a driver's emotion by investigating the relation between EEG, driving record, brightness, eye tracking (eye shut), and facial recognition. At this moment, we found the relation as shown in Figure 9. Unfortunately, we could not collect enough data to analyze the relation between these data and other data such as car operation (pedals, steering), heart rate, and body motion.

Firstly, we compared the facial expression estimation of fatigue by face recognition (Fig. 10) and the number of eyes

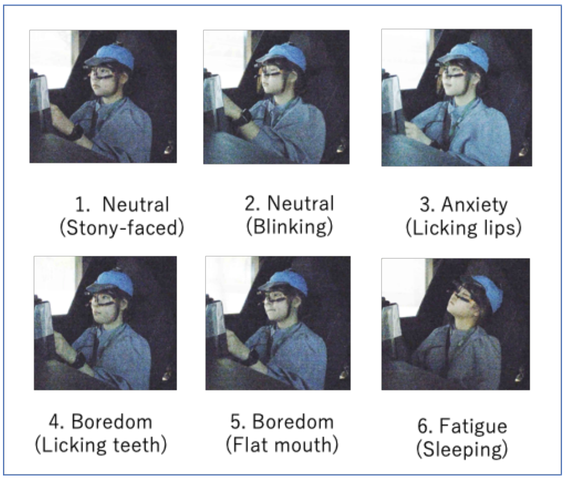


Fig.8 Six patterns from facial recognition

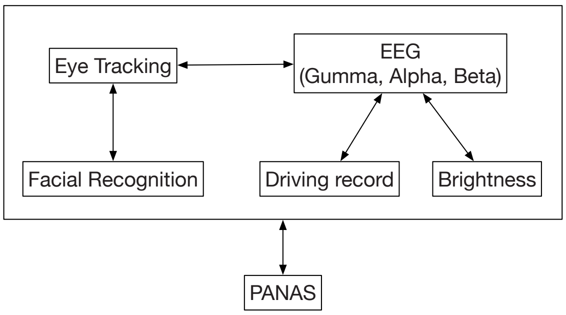


Fig. 9 Relation of information that was found from the experiment

shut (= sleepiness) by eye tracker (Fig. 11). The number of eye closures tends to increase with each round of testing. The same trend is observed on different experimental days and with different subjects. In contrast, there is a difference in the trend of facial expression estimation of fatigue even when the same

TABLE III
THE NUMBER OF FATIGUE AND EYE SHUT EVERY FIVE MINUTES.

	~05:00	~10:00	~15:00	~20:00	~25:00	~30:00	~35:00	~40:00	~45:00
Fatigue	1	5	1	1	4	5	3	10	9
Eye shut	2	2	4	0	3	5	6	13	0

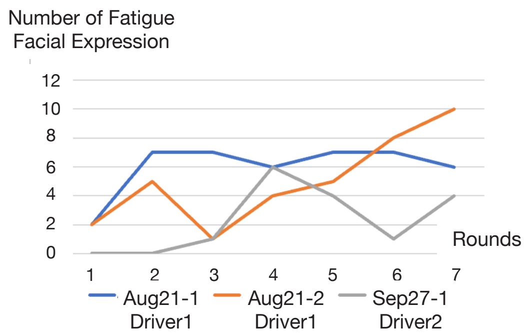


Fig.10 Number of fatigue from facial recognition (Paris, different driver)

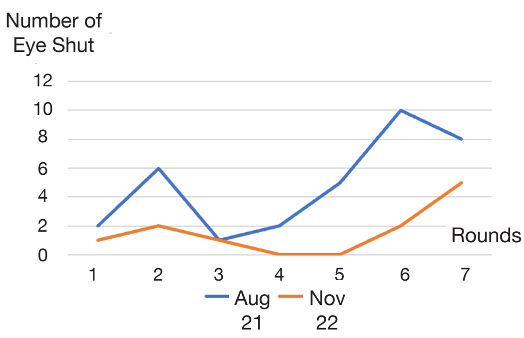


Fig.11 Number of eye shut (Paris, different driver)

subject is tested on the same day (Aug. 21-1 and Aug. 21-2 in Fig. 10). This result suggests that there is a limitation in finding fatigue estimation by face recognition.

For the lap time, as shown in Figure 5, each additional lap is getting longer. This result suggests a similar trend between eye tracking data and the number of laps (= driving time).

Then, we compared the EEG and other data. TGAM outputs raw data at 512 Hz and aggregated values of each brain wave band (Delta, Theta, Alpha, Beta, and Gamma) per second. However, it is difficult to read their interrelationships because EEG changes significantly from one second to the next. Therefore, we applied topological data analysis (TDA) [29] to analyze EEG data (Alpha, Beta, and Gamma). The basic technology of TDA is persistent homology. The following is an overview of persistent homology from [29].

"A key mathematical apparatus in TDA is *persistent homology*, which is an algebraic method for extracting robust topological information from data. To provide some intuition for the persistent homology, let us consider a typical way of constructing persistent homology from data points in a Euclidean space, assuming that the data lie on a sub- manifold. The aim is to make inference on the topology of the underlying manifold from finite data. We consider the r-balls (balls with radius r) to recover the topology of the manifold, as popularly employed in constructing an r-neighbor graph in many manifold learning algorithms. While it is expected that, with an appropriate choice of r, the r-ball model can represent the underlying topological structures of the manifold, it is also known that the result is sensitive to the choice of r. If r is too small, the union of r-balls consists simply of the disjoint r-balls. On the other hand, if r is too large, the union becomes a contractible space. *Persistent homology* [30] can consider *all* r simultaneously, and provides an algebraic ex- pression of topological properties together with their persistence over r. The persistent homology can be visualized in a compact form called a *persistence diagram* $D = \{(b_i , d_i) \in R^2 \mid i \in I, b_i \leq d_i\}$, and this paper focuses on persistence diagrams, since the contributions of this paper can be fully explained in terms of persistence diagrams. Every point $(b_i, d_i) \in D$, called a *generator* of the persistent homology, represents a topological property (e.g., connected components, rings, and cavities) which appears at X_{b_i} and dis- appears at X_{d_i} in the r-ball model. Then, the *persistence* $d_i - b_i$ of the generator shows the robustness of the topo- logical property under the radius parameter. "

We used HomeCloud [31], a tool for visualizing persistence; we created a 3D graph of Alpha, Beta, and Gamma for the Paris orbit on Aug.21, 2023 (2nd trial, by driver1) and the Paris orbit on Nov.22, 2023 (5th trial, by driver2). The graph of persistence generated from the 3D data of Alpha, Beta, and Gamma is shown in Figure 12. Most of the points lie on the X=Y line, but the points away from it represent the features of the data.

In Figure 13, the envelopes are added for the big-picture view of the points. It can be seen that the shape of the envelope for each lap is similar, even though two drivers with different driving skills and different schedules are driving on different dates. This result indicates that some changes may be occurring similarly for each lap. In the future, we plan to analyze the EEG movements related to fatigue by analyzing the EEG in more detail.

VII. CONCLUSION

This study aims to estimate driver emotion by using several data that we can acquire while driving a car to prevent car accidents. For that purpose, we developed a sensor network around a driving simulator using an EEG sensor, accelerometer, heart rate sensor, and eye tracker.

The paper's novel contribution is that the result displayed a relation between the driver's emotions, especially sleepy

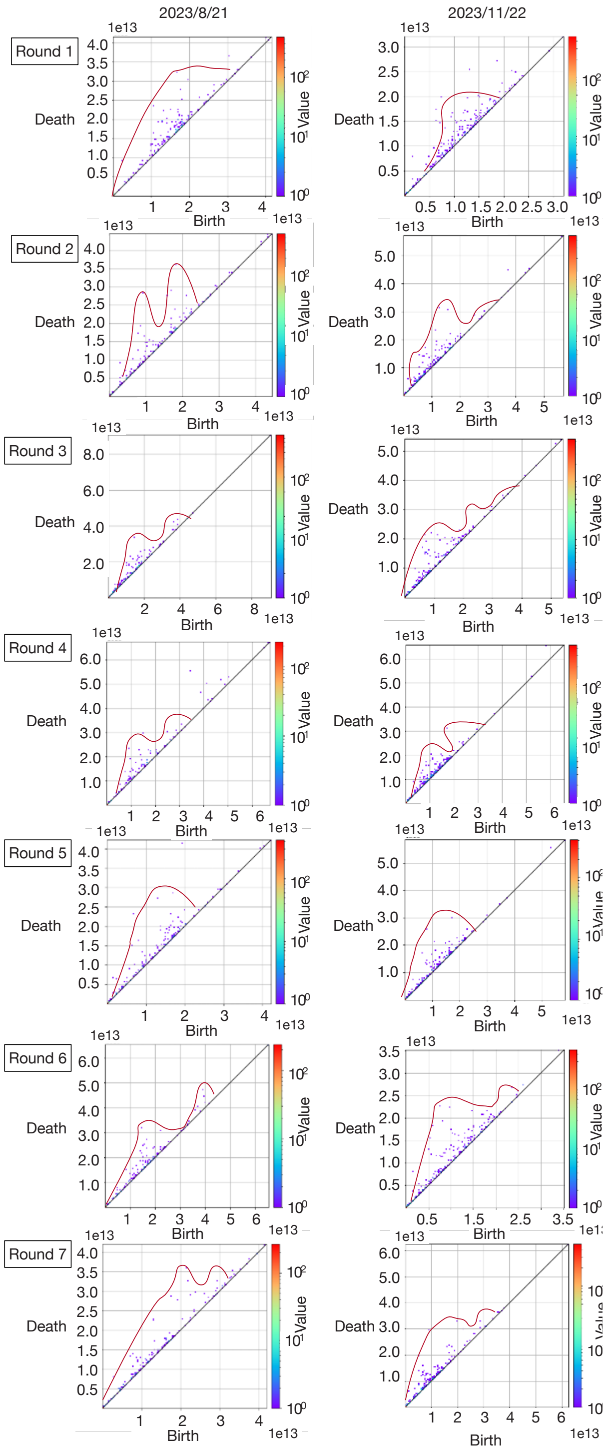


Fig. 12 Persistent homology graph of brain waves (Alpha, Beta, and Gamma)

conditions, driving speed, driving duration (=driving distance), and brain wave behavior. This result will make it possible to realize a safety-driving support technology.

However, at this moment, as motioned in section VI, our analysis was not yet sufficient because of the limited number of data. If we get enough data, we could understand the relation of information around a car to understand the driver's emotions.

However, this study showed the effectiveness of Cognitive Mobility, that is a part of Cognitive Infocommunication science.

As a further study, we need to do two things. The first is to analyze the data and information relation using new techniques such as persistent homology. The second is to acquire the Mental State Index, such as PANAS while trying to explain the mental state in language.

Our final target is to realize a system to detect the dangerous mental state related to car accidents and provide information to change the situation using a simple wearable sensor such as a smartwatch and CAN data.

Acknowledgment

The authors would like to express special thanks to Mr.Inagaki, Mr.Yuchi, Mr.Miyagawa, Mr.Sasuga, Mr.Kosuge, Ms.Tachibana, Mr.Gunji of Toyota Customizing & Development Co., Ltd. for allowing us to perform this research.

References

[1] M. Mahsa, D. Dale and K. Arshia. (2020). Electroencephalography (EEG) Technology Applications and Available Devices. *Applied Sciences*. DOI: 10.3390/app10217453.

[2] P. Baranyi and A. Csapo, "Definition and Synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp.67– 83, 2012, DOI: 10.12700/aph.

[3] P. Baranyi, A. Csapo, and G. Sallai, "Emergence of CogInfoCom in Science and Technology," in *Cognitive Infocommunications (CogInfoCom)*, Springer, Cham, 2015, DOI: 10.1007/978-3-319-19608-4_1.

[4] Zöldy, M., & Baranyi, P. (2021). Cognitive Mobility – CogMob. In *2021 12th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)* (pp. 000 915–000 919).

[5] Zöldy, M., & Baranyi, P. (2022). The Cognitive Mobility Concept. *Infocommunications Journal*, 2023/01, pp. 35–40

[6] Gianluca Borghini, Laura Astolfi, Giovanni Vecchiato, Donatella Mattia, Fabio Babiloni, Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness, *Neuroscience & Biobehavioral Reviews*, Volume 44, 2014, Pages 58–75, ISSN 0149-7634, DOI: 10.1016/j.neubiorev.2012.10.003.

[7] Takanori CHIHARA, Fumihiro KOBAYASHI and Jiro SAKAMOTO, Estimation of mental workload during automobile driving based on eye-movement measurement with a visible light camera, *Transactions of the JSME (in Japanese)*, Vol.86, No.881, 2020

[8] Edmund Wascher, Emad Alyan, Melanie Karthaus, Stephan Getzmann, Stefan Arnau, Julian Elias Reiser, Tracking drivers' minds: Continuous evaluation of mental load and cognitive processing in a realistic driving simulator scenario by means of the EEG, Heliyon, Volume 9, Issue 7, 2023, e17904, ISSN 2405-8440, DOI: 10.1016/j.heliyon.2023.e17904.

[9] Magliacano A, Fiorenza S, Estraneo A, Trojano L. Eye blink rate increases as a function of cognitive load during an auditory oddball paradigm. *Neurosci Lett*. 2020 Sep 25;736:135293. Epub 2020 Aug 6. PMID: 32771601. DOI: 10.1016/j.neulet.2020.135293.

[10] C. Hexuan, L. Yingzi and M. Ronald. (2007). Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators. DSC North America – Iowa City – September North America – Iowa City – September.
[11] M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
[12] A. Kolli, A. Fasih, F. A. Machot and K. Kyamakya, "Non-intrusive car driver's emotion recognition using thermal camera," *Proceedings of the Joint INDS'11 & ISTET'11*, 2011, pp. 1–5.
[13] X. Huafei, L. Wenbo, Z. Guanzhong, W. Yingzhang, X. Jiyong, Z. Juncheng, L. Chengmou and G. Gang. (2022). On-Road Driver Emotion Recognition Using Facial Expression. *Applied Sciences*. 12. 807. DOI: 10.3390/app12020807.
[14] M. Patil and S. Veni, "Driver Emotion Recognition for Enhancement of Human Machine Interface in Vehicles," *2019 International Conference on Communication and Signal Processing (ICCSP)*, 2019, pp. 0420–0424.
[15] Zaman, K., Sun, Z., Shah, S.M., Shoaib, M., Pei, L. and Hussain, A., 2022. Driver Emotions Recognition Based on Improved Faster R-CNN and Neural Architectural Search Network. *Symmetry*, 14(4), p. 687.
[16] Priyanka A. Abhang, Bharti W. Gawali, Suresh C. Mehrotra, Chapter 2 – Technological Basics of EEG Recording and Operation of Apparatus, Editor(s): Priyanka A. Abhang, Bharti W. Gawali, Suresh C. Mehrotra, Introduction to EEG- and Speech-Based Emotion Recognition, Academic Press, 2016, Pages 19–50, ISBN 9780128044902, DOI: 10.1016/B978-0-12-804490-2.00002-6.
[17] Mindsall Inc., https://mindsall.com
[18] M5Stack Core2 ESP32 IoT Development Kit, https://shop.m5stack.com/products/m5stack-core2-esp32-iot-development-kit-v1-1
[19] TGAM datasheet, https://cdn.hackaday.io/files/11146476870464/TGAM%20Datasheet.pdf
[20] Pupil Core, https://pupil-labs.com/products/core/
[21] Bluetooth Technology Overview, https://www.bluetooth.com/learn-about-bluetooth/tech-overview/
[22] Controller Area Network – an overview, https://www.sciencedirect.com/topics/engineering/controller-area-network
[23] SUMO, https://eclipse.dev/sumo/
[24] David Watson;Lee Anna Clark(1999), “The PANAS-X: Manual for the Positive and Negative Affect Schedule – Expanded Form”
[25] Crane, E., Gross, M. (2007). Motion Capture and Emotion: Affect Detection in Whole Body Movement. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds) Affective Computing and Intelligent Interaction. ACII 2007. *Lecture Notes in Computer Science*, vol 4738. Springer, Berlin, Heidelberg.
[26] MediaPipe Team, Face Mesh. MediaPipe, 2022 [online].
[27] Opencv-python 4.6, 2022 [online]. https://pypi.org/project/opencv-python/.
[28] J. Luo, H. Yoshimoto, Y. Okaniwa, Y. Hiramatsu, A. Ito and M. Hasegawa, "Emotion monitoring sensor network using a drive recorder," *2023 IEEE 15th International Symposium on Autonomous Decentralized System (ISADS)*, Mexico City, Mexico, 2023, pp. 1–8, DOI: 10.1109/ISADS56919.2023.10092139.
[29] Genki Kusano, Yasuaki Hiraoka, Kenji Fukumizu, Persistence weighted Gaussian kernel for topological data analysis, *Proceedings of The 33rd International Conference on Machine Learning, PMLR* 48:2004-2013, 2016. *ArXiv*:1601.01741
[30] Edelsbrunner, H., Letscher, D., and Zomorodian, A. Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4):511–533, 2002.
[31] HomCloud, https://homcloud.dev/index.en.html

Jinshan Luo received B.A. degrees in science from Chengdu University in 2016 and M.A. degrees in engineering from Utsunomiya University in 2021, respectively. He is now a Doctor's Course student at a department of engineering at Utsunomiya University since 2022. He is also a researcher for Toyota Customize and Development Co., Ltd. since 2021. His current research interests include mental state estimation, data science, image recognition, sensor network, and autonomous driving.

Yuki Okaniwa is a student of Faculty of Economics of Chuo University.

Natsuno Yamazaki is a student of Faculty of Economics of Chuo University.

Yuko Hiramatsu received the B.A. and M.A. degrees in Literature from Sophia University in 1979 and 1985, respectively. She has been on the secretariat staff of the “Study Group of School Trip” since 2006. She was an adviser for the utilization of ICT in Okinawa in the subsidy project of the Ministry of Internal Affairs and Communications in Japan from 2007 to 2009. She belongs to the Faculty of Economics at Chuo University as a lecturer since 2009 and a specially appointed professor from 2023. Her current research interests are the linguistic landscape and ICT literacy for primary education.

Madoka Hasegawa received B.S., M.S., and Ph.D. degrees in engineering from Utsunomiya University in 1994, 1996, and 1999, respectively. She is now a professor at the school of engineering, Utsunomiya University. Her current research interests include image processing, image coding, digital watermarks, and usable security. She is a senior member of IEICE (the Institute of Electronics Information and Communication Engineers, Japan) and a member of IEEE.

Atsushi Ito received B.S. and M.S. degrees from Nagoya University in 1981 and 1983, respectively. He also received a Ph.D. degree from Hiroshima City University in 2007. From 1983 to 2014, he was with the Research and Development Laboratories of KDDI Corporation, and he was a Professor at the Graduate School of Engineering of Utsunomiya University from 2014 to 2020. He has been a Professor at the Faculty of Economics of Chuo University since 2020. During 1991-1992, he was a visiting scholar at the Center for the Study of Language and Information (CSLI) at Stanford University. His current research interests include ad hoc networking, biosignal applications, and ICT-based agriculture.

Special Issue
of the Infocommunication Journal

Developing sustainable logistic strategies in the context of cognitive biases

Péter Földesi[1], and Eszter Sós[2]*

Abstract—**Cognitive biases often occur even in the decision-making process of highly qualified company managers due to the drive for efficiency and time pressure in operations. At the same time, there are also long-term strategic decisions where time pressure is no longer a factor, and yet cognitive bias appears, which has to be considered properly. In strategic issues, decision- makers tend to see their wishes and desires rather than the objective reality. The proposed system of fuzzy indicators based on technical and objective data supports decision-making between logistics strategies by mitigating cognitive biases, which is extremely important in the logistics field, where the decisions have to be made partly based on subjective, vague, or uncertain parameters.**

Index Terms—**logistics systems, cognitive bias, Push and Pull strategy, fuzzy description**

I. Introduction and Background

Cognitive infocommunication aims to create complex perceptual computing systems that effectively support human-machine communication [38]. The development of new methods, mathematical modeling, learning techniques, and related behavioral research will also help to better understand perceptual and cognitive brain processes [39]. For human-machine communication to be effective in the course of logistical strategies, it is necessary to identify and avoid cognitive biases in decision-making, so that its evaluation is not only based on intuition and subjective judgment alone. Data and pre-processing, standardization, and quantification are necessary to avoid associated biases.

A cognitive bias is a systematic deviation from rationality and logical, reasonable thinking and behavior. Cognitive biases are phenomena that influence thinking on experiences, intuition, and perceived things [1], thereby turning an objective decision into a subjective interest system [2].

Logistics, as a specialized field, typically requires decision-making at a daily level for the staff implementing logistics processes. On the other hand, logistics is often not seen as an independent field of expertise. As a consequence, the importance and necessity of the processes concerned are not assessed in sufficient depth, and decisions are based on intuition rather than rational decision-making. The time pressure characteristic of the logistics field [3] also significantly affects cognitive biases. Due to the particularities of logistics, there are departments where self-interest distortion [4,5] appears among cognitive biases.

People's thinking works together with simplifications, and cognitive biases [6]. For this reason, cognitive bias often occurs at the decision-making level in objective interest systems, which can be interpreted in several ways [7]. A systematic deviation from economic rationality in a company's decision-making model is called bias [8]. Psychologist Gary Klein [9], - who studied intuition in a scientific context at length and analyzed its effects in decision-making situations - named it a recognition-based decision model [9]. He concluded that intuition-based decisions only help managerial decisions in a predictable environment, similar to what has already been experienced countless times. So, decisions based on intuition are only acceptable if they are based on real experience [10].

The starting point for the appearance of cognitive bias is always a situation where a person responsible for decision-making receives information that they must incorporate into their decision-making mechanism [11]. They try to support the decision-making processes objectively, by examining facts and data [12]. As a result of unconscious prejudices, "beliefs", and expected results, the examination of facts and data becomes subjective, and decision-making takes place without self-checking [13].

Logistics decision-making takes place at different levels: it can distinguish between tactical, operational, and strategic decisions, which can be grouped according to a time horizon into short-, medium- and long-term decisions. In this paper, the 5 categories of cognitive biases will be described [14] and identified in the course of corporate decision-making and how and at what level these biases appear in the course of logistics decision-making will be shown.

The choice between Push and Pull systems are examined [15] as well as providing insight into the cognitive biases that appear during decision-making related to these systems. As it is not only determined by whether the production takes place to customer demand or stock, decision-making is often influenced by the cognitive biases that arise from fears about stock issues. First, where the names of the Push and Pull systems come from is mapped [16], and then the history of their development is presented [17]. After clarifying the concepts, it is explained which cognitive biases can appear during the logistics-related decision-making mechanism. Further, the limit will be examined at which a product can be reasonably defined to have

[1] Department of Logistics and Forwarding, Széchenyi István University, Győr, Hungary and Hungarian Research Network, Budapest, Hungary
[2] Department of Logistics and Forwarding, Széchenyi István University, Győr, Hungary. (E-mail: sos.eszter@ga.sze.hu)

DOI: 10.36244/ICJ.2024.5.13

a Push or Pull logistics strategy associated with it. For this purpose, a fuzzy measure is proposed that can be used to clearly identify the logistics strategy for the given product and the company's production.

The choice between Push and Pull logistics strategy for sustainability

One of the big trade-offs in Push-Pull is inventory versus delivery cost. This gives rise to logistical strategy decisions, during which cognitive biases may emerge as a significant deciding factor. The cognitive biases that emerge during the logistics decision-making process are introduced in Chapter 3, and the Push and Pull logistics strategies are discussed in detail in Chapter 4.

In this article, an analysis was carried out on the ScienceDirect and Scopus site on whether sustainability as a goal is reflected in the choice between Push and Pull logistics strategies. The aim of this publication is to present a methodology; therefore, the literature analysis is limited to these two databases.

An analysis was conducted on ScienceDirect and Scopus, first examining all publications, then narrowing it down to the last 5 years (2019-2024) and the 15 years before that (2004-2018), where the keywords in Table 1 and their contexts were included in the filtering: "push pull" AND (logistics OR "supply chain management") AND sustainability on ScienceDirect on the 10th of April 2024.

TABLE I.
NUMBER OF RELEVANT PAPERS. KEYWORDS: "PUSH PULL" AND (LOGISTICS OR "SUPPLY CHAIN MANAGEMENT") AND SUSTAINABILITY

PERIOD	KEYWORDS / RESULT SCIENCEDIRECT	NUMBER OF PUBLICATIONS	
		Science Direct	Scopus
2000-2024	Summary	326	1065
2019-2024	Last 5 years	159	834
2004-2018	15 years before the last 5 years	141	231

Table 1 shows that the search did not yield a large number of results, suggesting that there is currently little research on the relationship between logistics strategies and sustainability.

To examine the keywords of the publications, using the VosViewer software, which helps us analyze the relationships between keywords in the publications, showing the direction of the articles written in recent years in the context of Push and Pull logistics strategies and sustainability

Visualizing all the results of ScienceDirect (Appendix 1), it can be seen that the supply chain appears directly next to sustainability from 2018, and Industry 4.0 appears in 2021-2022, in addition to the circular economy, strategy, blockchain, technology adoption, and flexibility keywords. This led to the conclusion that professionals have recognized the need to develop long-term logistics strategies to achieve sustainability, yet the publications that have appeared have not focused on this aspect.

Based on the keywords, a visualization was made from the Scopus database, where the minimum number of occurrences was set to 20 due to the large number of keywords. In Appendix 2 it is already clear that from 2019-2020, terms related to environmental will also appear: environmental management, environmental technology, environmental regulations. Hence, environmental measures related to sustainability have become more prominent in the last 5 years. Appendix 2 also illustrates that in 2022, the keywords consumption behaviour and human will already appear, highlighting that more publications have already examined the human perspective in terms of logistics strategies and sustainability.

This paper aims to highlight the need to examine the Push and Pull systems and the cognitive biases in the choice of strategies in terms of sustainability, as the logistics strategy is fundamental to this.

II. LOGISTICAL ASPECTS OF COGNITIVE BIASES APPEARING IN CORPORATE DECISION-MAKING

In logistics, sustainability is primarily a strategic decision [18, 19]. Inappropriate logistics sub-processes damage the environment by purchasing unnecessary equipment, parts or packaging [20]. If logistics strategies are not internally coherent, this has an impact on sustainability. The presence of cognitive biases can be recognized in almost all areas of logistics. For their appearance to become clearly identifiable, cognitive biases are first described in the following.

Cognitive psychology basically defines two thinking systems. One system is characterized by conscious and processed thoughts [7], this is what is called "rational thinking" [21], which is not dealt with in this publication. The other thinking system is characterized by automatic and intuitive thoughts. The presence of experiences, prejudices, and assumptions causes cognitive biases in the system of thinking processes, thereby simplifying decision-making situations [22].

From the point of view of corporate decision-making, Olivier Sibony [14] classified cognitive biases into five main categories, within which 23 different types (see Appendix 3) were distinguished. Logistical aspects of cognitive biases appearing during corporate decision-making are as follows:

Pattern-recognition biases: Arise when a company tries to follow the example of a successful person by incorporating the same decisions into corporate strategies, but these decisions are not always appropriate for the company, which has a completely different corporate culture and product range [14].

Action-oriented biases: Refers to the cognitive biases that appear in the actions, which usually result from overly optimistic planning [23].

Cognitive inertia: When a process does not start due to certain facts and data. Loss aversion is one of the most powerful cognitive biases, which hinders change and the development opportunity that comes with it [24]. Loss aversion is of great importance in the subfields of logistics [25].

Self-interest biases: In some cases, for managers, the most important thing is not the company's lost money, but the loss of prestige resulting from their failure. Self-interest bias, also known as limited ethics, refers to cognitive biases that cause decent people to unknowingly show unethical behavior [4, 26].

Group biases: The cognitive biases created in the group mean that one individual's opinion has a significant influence on the decision-making structure of other people [27]. The effect of group bias is significant for the preparation of logistics strategies since logistics is often still negatively evaluated.

Based on the five categories, it can be clarified that the presence of cognitive biases in the decision-making mechanism related to logistics strategies can be definitely identified.

During the work of the persons responsible for the implementation of logistics processes in corporate decision-making, the logistics mindset often appears, casting logistics in a bad light, and causing in turn extra work and costs for the company. In practice, this means that the work of the people responsible for the implementation of logistics processes is treated as unwelcome and an unnecessary cost, so the managers of smaller companies tend to follow this way of thinking, both in terms of their work and their financial appreciation.

III. Concepts of Push and Pull strategies

The terms Push and Pull first appeared in Richard J. Schonberger's book in 1982 [16], comparing the Western-oriented "Push production system", which is based on the design philosophy of production resources and material resource planning; and the Japanese "Pull production system", which included the control technique based on Kanban logic together with the expectations of the Just In Time concept [36].

Nowadays, the definition of Push and Pull is already defined [15, 17], but professionals in the industry still often use the concepts of Push and Pull incorrectly, and consequently make bad strategic decisions. One of the big trade-offs in Push-Pull is inventory versus delivery cost. This gives rise to logistical strategy decisions, during which cognitive biases may emerge as a significant deciding factor.

In the case of a Push system, preliminary demand surveys are carried out, based on which the production program is prepared, the raw material is procured in the appropriate schedule, and then the production program is executed (Material Requirements Planning and Manufacturing Resource Planning MRP II).

In the case of Pull systems, production is always initiated by customer demand. In this way, the minimum stock level of the finished product can be ensured, but on the other hand, it means a longer lead time for the customer. The Just In Time (JIT) system enjoys great popularity among companies. However, its name causes a misunderstanding in the common language. From whose point of view is the product just in time? Undoubtedly from the point of view of the inventory as, due to the lead time, the customer always ends up waiting.

There is a trade-off in logistics between inventory or delivery (mobility). It connects to logistics strategies if the product is Push, more stock is needed and less transport, so you have less mobility, which means you have lower CO_2 emissions and less noise pollution. Conversely, if the Pull start strategy is investigated, there will be a higher number of deliveries associated with the freight transport, as the delivery of inventory is determined by customer demand. However, cognitive mobility [40] can lead to a deviation from what seems to be a good solution based on objective calculation.

The history of the development of Push and Pull systems dates back to the appearance of Material Requirements Planning [28], which enables the planning of material requirements for production and procurement [30]. The starting point was the number of final products defined in the production program, next the bill of materials was determined based on the material requirement, and then the gross component and raw material requirement [31].

The Material Requirements Planning (MRP) system was developed by Joseph Orlicky for the Toyota Manufacturing Program in 1964 [32]. Simultaneously, Black & Decker was the first company to use MRP. By 1975, MRP was implemented in 700 companies, and Joseph Orlicky's book Material Requirements Planning [32] was published in the same year. In 1983, Oliver Wight put the master schedule, rough capacity planning, capacity requirement planning, and other concepts into the classic MRP, thereby creating the basic idea of Manufacturing Resource Planning (that is MRP II) [33].

The Pull logistics strategy started with the Kanban system, which was introduced in the 1940s in supermarkets [32]. The order was determined based on the seller's inventory [15]. They only ordered more when the stock of the item was significantly reduced, thus optimizing the flow between the supermarket and the consumer. Toyota engineers noticed this method and, led by Taichi Ohno, investigated how it could be applied to work processes in the industry [33]. The Kanban initiates an action to replace the quantity consumed, so it is assigned to each production lot within the Just In Time (JIT) system [34,35]. To make Kanban effective, cycle time must be assigned [36].

Cognitive biases appear during decision-making between Push and Pull systems – case study

When a new product is introduced, the decision regarding the logistics strategy associated with it is generally made rationally and based on sound arguments. However, some concerns often lead to cognitive biases during the decision-making process related to Push and Pull strategies [17, 28], such as fear of the Bullwhip effect [29], lack of supplies, etc. An improperly chosen logistics strategy can cause supply disruptions and is also reflected in the company's processes. For example, the size of the warehouse is not only determined by how many products the company sells per month, but it is also significantly influenced by whether the manufactured products are made in a Push or Pull system. A product made entirely in the Pull system is not stocked, or only for a very short time. Products manufactured in the Push system are produced in large quantities [15], so they require a larger storage capacity.

Furthermore, it is a common problem for companies to use a Pull strategy until the managers are faced with the fact that the logistics processes are not working well. When investigating the reasons, it turns out that the Pull strategy was chosen because the manager who made the decision had previously worked for a company where the products were associated with the Pull logistics strategy. The decision is therefore accepted without any examination of the external circumstances. Considering Sibony's classification [14] (see Appendix 3) the presence of Overconfidence bias can already be clearly identified during decision-making (see Table.2).

TABLE II.
CROSS-CORRELATION OF BIASES (OWN COMPILATION BASED ON [14])

		Own knowledge	
		Underestimates	Overestimates
Position (power)	Underestimates	1 - Loss Aversion	2 - Experience bias
	Overestimates	3 - Halo effect	4 - Overconfidence

A cognitive bias appeared in the manager's decision-making mechanism [10], which, based on experience so far [11], led to the conclusion that the logistics strategy being implemented made the previous company successful, even though the product requires a completely different logistics environment. As a result of the appearance of the Halo effect [24], the appropriate logistics strategy was not implemented. The decision is partly based on existing work experience, so the presence of Experience bias [1] can be identified. After that, Status Quo [4] is a common phenomenon, when the management sticks to the originally formed decision. When it turns out that the right strategy has not been chosen for the given product, Loss aversion, Uncertainty aversion, and Hindsight bias [4; 14], which often arise during further decisions, appear as well.

IV. DEMARCATION OF PUSH AND PULL LOGISTICS STRATEGY

Push and Pull systems are determined based on the needs related to the product [15], but in many cases during production, it cannot be clarified whether a specific product is manufactured in a Push or Pull system. Fig. 1/a shows that the Completely Push System is when first the forecast (F) is made, the raw material is purchased, and only then does production begin, with the finished product then being sold by the vendors. In contrast, Fig. 1/b shows that only after the customer's order do the raw material procurement and production start. Then, when the order is fulfilled, the customer receives the product. In the Hybrid Push/Pull system shown in Fig. 1/c, the first half of the system behaves as a Push. The Material Decoupling point appears at the same time as the semi-finished stock, after which the product starts to behave as a Pull.

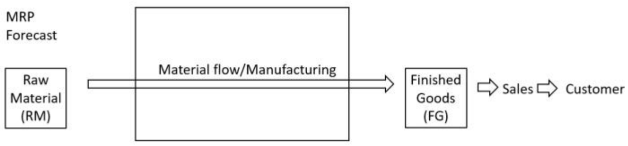


Fig.1/a. Completely Push System (MTS – Make-to-stock)

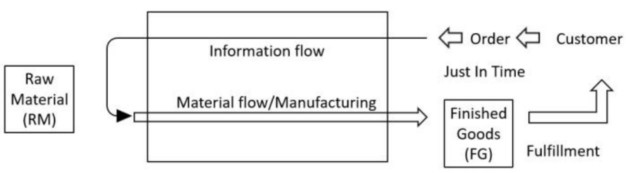


Fig.1/b. Completely Pull System (MTO – Make-to-order)

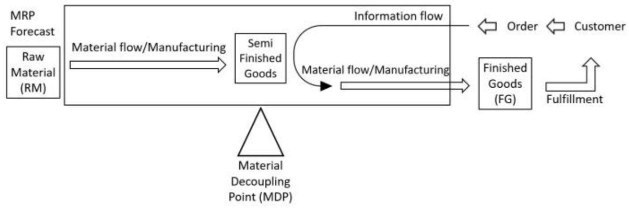


Fig.1/c. Hybrid Push/Pull System with Material Decoupling Point

V. FUZZY ASSESSMENT FOR THE DEMARCATION

An additional question is what characteristics can be used to determine whether it is Push or Pull. The first approach can be based on how many phases, for how long, and at what cost the material flow will proceed. For example, Fig. 2. shows a sequential production line

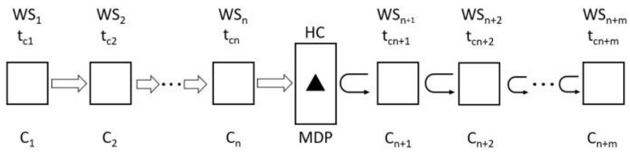


Fig.2. A sequential production line

Legend for sequential production line

Symbol	Meaning
WS_i =	The i^{th} Work Station ($i= 1...n+m$)
t_{ci} =	Cycle Time at WS_i ($i= 1...n+m$)
C_i =	Production Cost at WS_i ($i= 1...n+m$)
HC =	Holding Cost ($_{RM}$- raw material, $_{FG}$- finished goods)

The raw material arrives by WS_n, in the Push system. After that, it goes to the warehouse, and when an order comes, it will be pulled in small Kanban circles until WS_{n+m} (Fig.2). The MDP coincides with the entry into storage. There are $n+m$ workstations in total. Of these, n units are Push-based and m are Pull-based, the proportions of which are determined:

$$\text{Push: } \frac{n}{n+m} \qquad (1)$$

$$\text{Pull: } \frac{m}{n+m} \qquad (2)$$

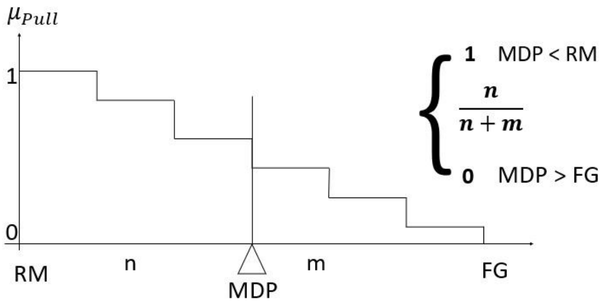


Fig.3. Fuzzy membership function of being pull calculated by (2)

This relationship considers only the used capacity in discrete numbers and can be used if the cycle times are nearly constant as shown in Fig. 3. For the representation of the actual status of the production concerning pull and/or push strategies fuzzy sets are introduced in which the membership values of being "Pull" are calculated at different accuracy, depending on what features of the production are taken into account (see Eq. (4), Eq. (6) and Eq. (8))

If the workload is not evenly distributed, it is necessary to consider the actual time spent in production (see Eq. (4)):

$$\text{Push: } \frac{\sum_{i=1}^{n} t_{ci}}{\sum_{i=1}^{n} t_{ci}+\sum_{j=n+1}^{n+m} t_{cj}} \quad (3)$$

$$\text{Pull: } \frac{\sum_{j=n+1}^{n+m} t_{cj}}{\sum_{i=1}^{n} t_{ci}+\sum_{j=n+1}^{n+m} t_{cj}} \quad (4)$$

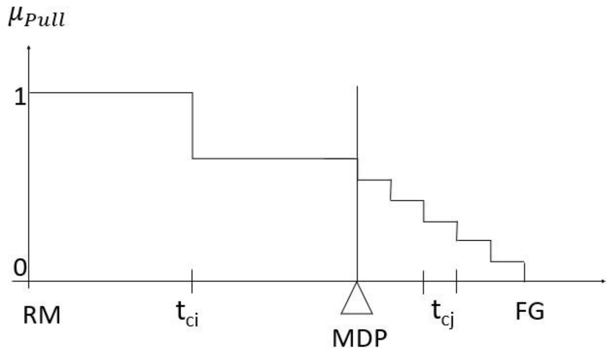


Fig.4. Fuzzy membership function of being pull calculated by (4)

By comparing Fig.3. and Fig.4. it is clear that the position of MDP alone is not enough to assess the “pullness” of the system. Without considering the cycle times judgement will be biased.

However, it does not yet include the complexity of the work process, for example, the processing may be different, such as rust protection, roughing, or fine machining. Hence a cost factor also must be included to determine the hourly price of the machine. In this case, C_i and C_j are introduced to represent the costs at each workstation (see Eq. (5) and Eq. (6)):

$$\text{Push: } \frac{\sum_{i=1}^{n} t_{ci}\cdot C_i}{\sum_{i=1}^{n} t_{ci}\cdot C_i+\sum_{j=n+1}^{n+m} t_{cj}\cdot C_j} \quad (5)$$

$$\text{Pull: } \frac{\sum_{j=n+1}^{n+m} t_{cj}\cdot C_j}{\sum_{i=1}^{n} t_{ci}\cdot C_i+\sum_{j=n+1}^{n+m} t_{cj}\cdot C_j} \quad (6)$$

C_i and C_j represent machine cost only, and there are further expenses to be considered. The cost of storage – in this case, Holding Cost *(HC)* -, is added to this (see Eq. (7) and Eq. (8)), where *RM* stands for raw material and *FG* stands for finished goods:

$$\text{Push: } \frac{HC+\sum_{i=1}^{n} t_{ci}\cdot C_i}{HC_{RM}+\sum_{i=1}^{n} t_{ci}\cdot C_i+\sum_{j=n+1}^{n+m} t_{cj}\cdot C_j} \quad (7)$$

$$\text{Pull: } \frac{\sum_{j=n+1}^{n+m} t_{cj}\cdot C_j}{HC_{FG}+\sum_{i=1}^{n} t_{ci}\cdot C_i+\sum_{j=n+1}^{n+m} t_{cj}\cdot C_j} \quad (8)$$

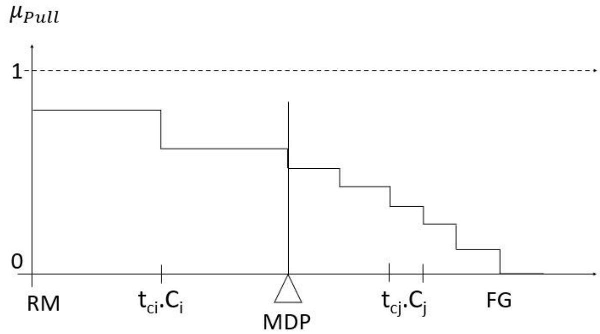


Fig.5. Fuzzy membership function of pull being calculated by (8)

If it contains HC_{RM}, it shows that the product is more of a Push product, since the Pull product is made for a specific period based on customer demand. So, the higher the *HC*, the more the system can be considered Push.

Important remark: Despite an apparently full pull system at first glance, when finished goods are stored for different reasons, for example, to consolidate the cargo and save cost in the delivery phase, in (Eq.8) HC_{FG} is not zero, the maximum value cannot reach *1*, and the sum of membership values of being push and being pull is also less than *1*. In this case, type2 fuzzy sets have to be applied that can describe the double uncertainty of the phenomena (see Fig. 5.).

The presented system of fuzzy representation is vital since - as it is demonstrated- even relatively simple situations can be very confusing and mislead the management. The proposed method can be used to objectively determine which logistic strategy is required for a given product, thus reducing the effect of Overconfidence bias in the process of decision-making.

The problem is well illustrated by the mask shortage that occurred during the COVID-19 pandemic. Mask production was a push system before the COVID-19 pandemic. A forecast was used to determine when and how many masks would be sold, and these forecast quantities were produced and then sold. With the onset of the COVID-19 pandemic, the forecast, based on decades of experience, was wiped out. Suddenly, production became a pull system, because when the raw material came in, they already knew which customer the particular quantity of masks they were producing belonged to. This publication demonstrates that the choice between Push and Pull systems is very confusing at times, as there are times when the Push logistics system is the right one, and then, in response to an unexpected situation, the product requires a Pull logistics system.

Accordingly, the system of equations described above can determine, for a given product or the company as a whole, whether the production uses a Push or Pull logistics strategy. Consequently, the decision-influencing effects of cognitive biases in the choice of logistics strategies can be avoided.

VI. Summary

Logistics is a frenetic field, one of the characteristics of which is frequent decision-making during the work process, accompanied all too often by time pressure. The importance of the logistics field was also brought to the fore by the COVID pandemic when most of society realized that if logistics systems do not work well, then the "supply of the world" does not work either. The disruption of supply chains also "overruled" the use of the previously popular Just In Time system, so the importance of choosing logistics strategies also increased.

This paper first presented an analysis of Push and Pull logistics systems and sustainability in terms of published papers. Despite the importance of addressing the impact of logistics strategies on sustainability, it was found that there are relatively few publications on the subject. After describing the emergence of cognitive biases, the 5 main categories of cognitive biases were introduced in corporate decision-making. For each main category, an example appearing in a logistics specialty was added to clarify the appearance of the types of corporate decision-making in the course of logistics decision-making. After describing the antecedents and history of the development of logistics strategies, the cognitive biases that appear during the choice between Push and Pull systems are illustrated through a concrete example. To clarify the choice between the Push and Pull strategy, three case studies were presented as well as a system of mathematical equations that support decision-making related to logistics strategies and help avoid the appearance of cognitive biases.

VII. Conclusion

This publication points out that the cognitive biases that appear during logistics decision-making have a significant impact on the formation of the logistics environment and logistics processes, so their influence on decisions cannot be ignored.

In the field of logistics, sustainability is first and foremost a strategic choice. Inappropriate logistics subprocesses damage the environment through the purchase of unnecessary equipment, parts, or packaging. If logistics strategies are not internally coherent, this, in turn, has an impact on sustainability. Cognitive biases occur during the decision-making process in the internal coherence of logistics strategies. Furthermore, the system may become unsustainable if an inappropriate logistic strategy is chosen because of cognitive bias. Therefore, it is also of paramount importance to avoid the appearance of cognitive biases in the decision-making process of logistics strategies to achieve sustainability.

The most important task of logistics strategy is to strengthen, coordinate and manage the relationship between corporate strategy and the logistics function, so that logistics can actively contribute to the company's performance and success. In this publication, only the Push and Pull logistics strategy was dealt with, which may seem simple, but if the right strategy is not chosen, it will not be sustainable.

Often Pull is chosen because it is fashionable and people think that keeping high stock is inappropriate. Assuming everything could be managed in Push and Pull, does not weigh the choice between strategies appropriately. However, if the Pull strategy is chosen unnecessarily, it can lead to high environmental impact, transport costs, pollution, and congestion in traffic areas.

It may be commercially worthwhile for a company because the external costs are not factored into the business model; society pays the cost, but it is not sustainable. In the short term, it seems right, but in the long term, it causes significant damage.

In order to give a proper description of the actual situation and help the decision-making, quantified measures have to be applied, and the fuzzy membership functions are capable of handling the issue. To develop the equation system, a sequential production line was first visualized, and then the workstations for the Push and Pull systems were defined. The membership values of the "Pull" state were calculated with different accuracies depending on the characteristics of the production. In order to estimate the "Pull" status, cycle times, workstation costs and storage costs had to be taken into account. The system of equations developed was used to convert the result into a type2 fuzzy number.

The proposed system of fuzzy indicators based on technical and objective data supports decision-making between logistics strategies by mitigating cognitive biases.

The presented and discussed set of fuzzy indicators based on technical and objective data are able to point out the real nature of the outlined production system from logistics point of view. The actual values of the calculated type2 fuzzy numbers are representation of the practical operation, so cognitive biases can be considered, which is extremely important in the logistics field, where the decisions have to be made partly on the basis of subjective, vague or uncertain parameters.

References

[1] Sibony O., Lovallo D., Powell T. C. Behavioral Strategy and the Strategic Decision Architecture of the Firm. *California Management Review*. 2017. Vol 59, Issue 3. DOI: 10.1177/0008125617712256

[2] Erceg N., Galić Z., Bubić A. Normative responding on cognitive bias tasks: Some evidence for a weak rationality factor that is mostly explained by numeracy and actively open-minded thinking. *Intelligence*. 2022. Volume 90, 101619 DOI: 10.1016/j.intell.2021.101619

[3] Albishri D. Y., Sundarakani B., Gomisek B. An empirical study of relationships between goal alignment, centralized decision-making, commitment to networking and supply chain effectiveness using structural equation modeling. *International Journal of Logistics Research and Applications. A Leading Journal of Supply Chain Management*. 2020. Volume 23, 2020 – Issue 4.

[4] Kahneman D., Tversky A. Prospect Theory: An Analysis of Decision Under Risk. *Econometric Society. 1979*. pp. 263–292. DOI: 10.2307/1914185

[5] Nadarevic L., Schnuerch M., Stegemann M. J. Judging fast and slow: The truth effect does not increase under time-pressure conditions. *Judgment and Decision Making*. 2021. Vol. 16, No. 5, pp. 1234–1266.

[6] Baron J. *Thinking and Deciding, 4th edition*. Cambridge University Press. 2006. ISBN: 9780511840265

[7] Toneatto T. Cognitive therapy for problem gambling. *Cognitive and Behavioral Practice*. 2002. Volume 9, Issue 3, pp. 191–199. DOI: 10.1016/S1077-7229(02)80049-9

[8] Stanovich K. E., West R. F. On the Relative Independence of Thinking Biases and Cognitive Ability. *Journal of Personality and Social Psychology 94*. 2008. pp. 672–695. DOI: 10.1037/0022-3514.94.4.672

[9] Klein G. *Sources of Power: How People Make Decisions*. Cambridge, MIT Press. 1998. ISBN13 9780262534291

[10] Kahneman D., Klein G. Conditions for Intuitive Expertise: A Failure to Disagree. *American Psychologist 64*. 2009. pp. 515–526.

[11] Dobelli R. *The Art of Thinking Clearly*. Sceptre. 2013. ISBN: 9780062359803

[12] Kahneman D., Tversky A. Subjective probability: A judgment of representativeness. *Cognitive Psychology*. 1972. pp. 430–454.

[13] Trimmer P. C. Optimistic and realistic perspectives on cognitive biases. *Current Opinion in Behavioral Sciences*. 2016. pp. 37–43.

[14] Sibony O. *You're about to Make a Terrible Mistake: How Biases Distort Decision-Making and What You Can Do to Fight Them*. Little, Brown and Company. 2020. ISBN: 0316494984

[15] Puchkova A., Romancer J-L., McFarlane D. Balancing Push and Pull Strategies within the Production System. *IFAC-Papers OnLine*. 2016. Volume 49, Issue 2 pp. 66–71. DOI: 10.1016/j.ifacol.2016.03.012

[16] Schonberger R. J. *Japanese Manufacturing Techniques*. Macmillan, USA. 1982. ISBN: 9780029291009

[17] Minculete G., Olar P. "Push" and "Pull" systems in supply chain management. Correlative approaches in the military field. *Journal of Defense Resources Management*. 2016. Vol. 7, No. 2. pp. 165–172. ISSN: 2068-9403

[18] Golam K., Sanjoy K. P. and Syed M. A. (2022) Sustainable Assessment in Supply Chain and Infrastructure Management. *Sustainability. 14(11)*, 6787. DOI: 10.3390/su14116787

[19] Rajeev A., Rupesh K. Pati, Sidhartha S. Padhi, Kannan, G. Evolution of sustainability in supply chain management: A literature review. *Journal of Cleaner Production*. Vol. 162, pp. 299–314. DOI: 10.1016/j.jclepro.2017.05.026

[20] Sós E., Földesi P. Fuzzy QFD assessment of logistics coherence. *2022 IEEE 6th International Conference on Logistics Operations Management (GOL)*. 2022. pp. 1–6. DOI: 10.1109/GOL53975.2022.9820009

[21] Waerneryd K. E. *Stock-market psychology*. Northampton, Edward Elgar Publishing. 2001. ISBN: 978 1 84064 736 5

[22] Kahneman D. Thinking, *Fast and Slow*. Macmillan Publishers. 2011. ISBN: 9781429969352

[23] Moore D. A., Healy P. J. The Trouble with Overconfidence. *Psychological Review 115*. 2008. pp. 5002–5017. DOI: 10.1037/0033-295X.115.2.502

[24] Rosenzweig P. *The Halo Effect... and the Eight Other Business Delusions That Deceive Managers*. New York, Free Press. 2007. ISBN: 1471137163

[25] Földesi P., Botzheim J.: Modeling of loss aversion in solving fuzzy road transport traveling salesman problem using eugenic bacterial memetic algorithm. *Memetic Computing 2*. 2010. pp. 259–271.

[26] Janis I. L. Groupthink: *Psychological Studies of Policy Decision and Fiascoes*. Boston, Wadsworth. 1982. ISBN-10: 9780395317044

[27] Bazerman M. H., Loewenstein G., Moore D. A. Why Good Accountants Do Bad Audits. *Harvard Business Review 80(11)*. 2008. pp. 96–102, 134.

[28] Puchkova A., Romancer J. L., McFarlane D.: Balancing Push and Pull Strategies within the Production System. *IFAC-Papers OnLine*. 2016. Volume 49, Issue 2, pp. 66–71. DOI: 10.1016/j.ifacol.2016.03.012

[29] Lin J., Zhou L., Spiegler V. L. M., Naim M. M., Syntetos A.: Push or Pull? The impact of ordering policy choice on the dynamics of a hybrid closed-loop supply chain. *European Journal of Operational Research*. 2022. Volume 300, Issue 1. pp. 282–295. DOI: 10.1016/j.ejor.2021.10.031

[30] Buzacott J. A.: Queueing models of Kanban and MRP controlled production systems. *Engineering Costs and Production Economics*. 1989. Volume 17, Issues 1–4, pp. 3–20. DOI: 10.1016/0167-188X(89)90050-5

[31] Milne R. J., Mahapatra S., Wang C-T. Optimizing planned lead times for enhancing performance of MRP systems. *International Journal of Production Economics*. 2015. Volume 167, pp. 220–231. DOI: 10.1016/j.ijpe.2015.05.013

[32] Kiran D.R. Chapter 30 – Material requirement planning. *Production Planning and Control, A Comprehensive Approach*. 2019. pp. 429–439. DOI: 10.1016/B978-0-12-818364-9.00030-5

[33] Kiran D. R. Chapter 31 – Manufacturing resource planning (MRP II). *Production Planning and Control, A Comprehensive Approach*. 2019. pp. 441–455. DOI: 10.1016/B978-0-12-818364-9.00031-7

[34] Sudarsan D. and Sivakumar G. D. Implementation of kanban to improve efficiency of a manufacturing system subject to bottlenecks. *International Journal of Manufacturing Technology and Management*. 2020. pp. 467–480.

[35] Kozma D. and Varga P. Supporting Digital Supply Chains by IoT Frameworks: Collaboration, Control, Combination. *Infocommunications Journal*, Vol. XII, No 4, December 2020, pp. 22–32. DOI: 10.36244/ICJ.2020.4.4

[36] Kiran D.R. Chapter 26 – Just in time and kanban. *Production Planning and Control, A Comprehensive Approach*. 2019. pp. 369–379. DOI: 10.1016/B978-0-12-818364-9.00026-3

[37] Olhager J. The role of the customer order decoupling point in production and supply chain management. *Computers in Industry*. 2010. Volume 61, Issue 9, pp. 863–868.

[38] Baranyi P., Csapo A., & Sallai G. (2015). *Cognitive infocommunications (coginfocom)*. Springer Book

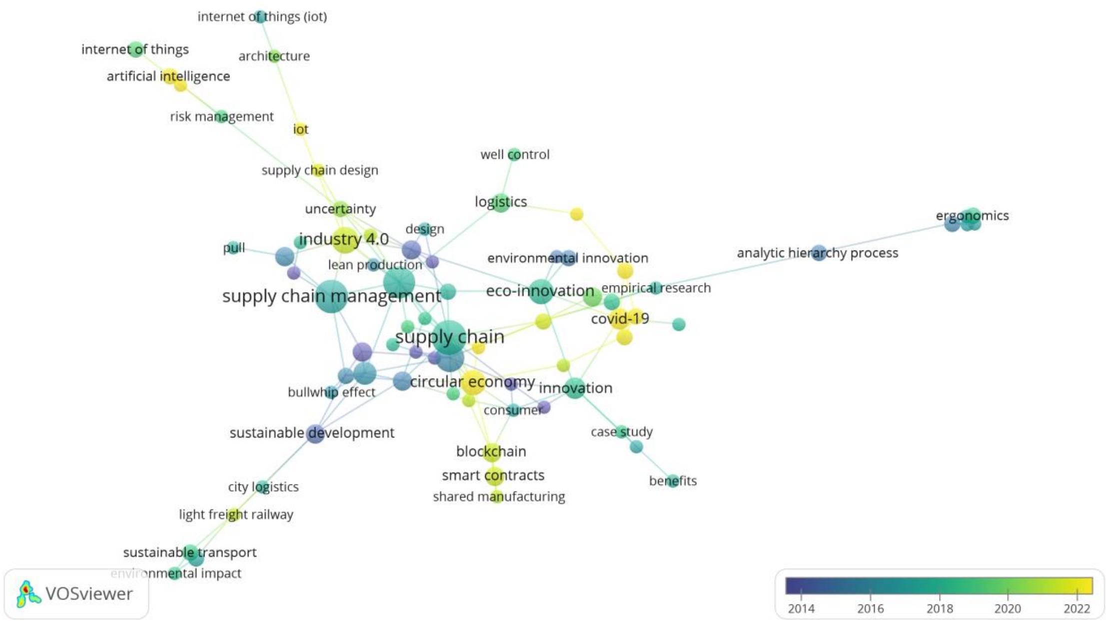
[39] Baranyi, P., & Csapó, Á. (2012). Definition and synergies of cognitive infocommunications. *Acta Polytechnica Hungarica*, 9(1), 67–83.

[40] Zöldy, M. & 39. Baranyi, P. (2023) The Cognitive Mobility Concept. *Infocommunications Journal*, 35–40.

Péter Földesi PhD was born in Budapest, Hungary in 1962. He graduated from the Budapest University of Technology and Economics as a Transport Engineer in 1986 and obtained PhD degree at the Hungarian Academy of Sciences in 1994. He worked for a large international haulage company as a marketing consultant in 1990, then joined the academic staff of Budapest University of Technology and Economics. In 1995 he moved to Győr (Hungary). He took a position at the Széchenyi István University, at the Department of Logistics and Forwarding, where he was promoted to Head of the Department in 2007. He was the Rector of the University between 2013 and 2021.

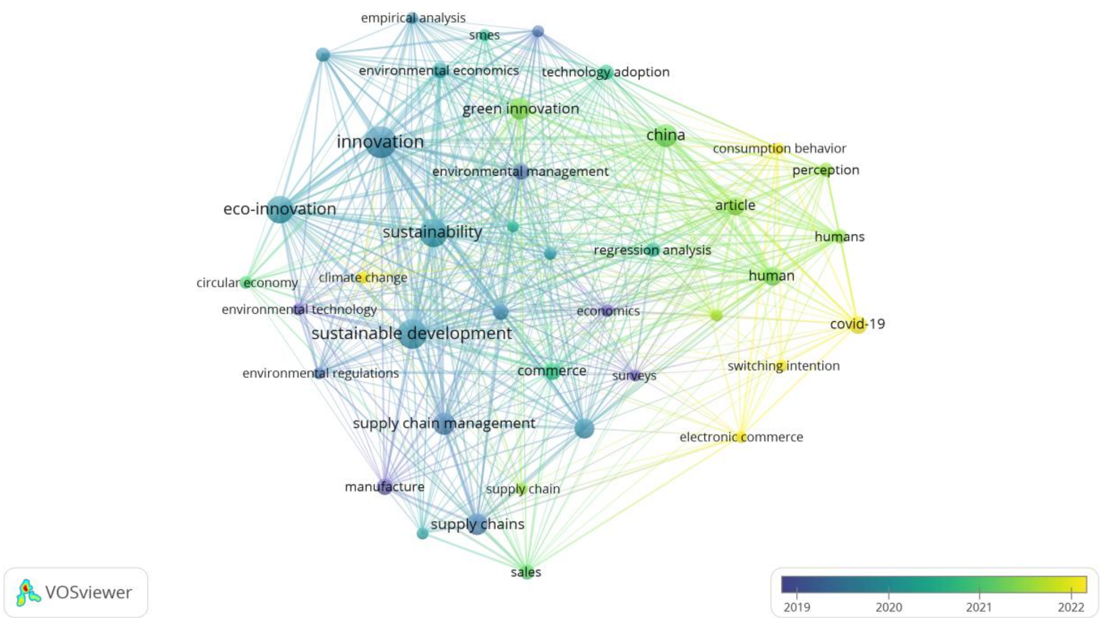
Eszter Sós was born in Bonyhád, Hungary in 1984. She graduated as a Logistics Engineer and is pursuing her PhD studies at the Doctoral School of Multidisciplinary Engineering Sciences at Széchenyi István University. She worked in the industry and logistics services for several years before starting her academic life. She is an Assistant Lecturer at Széchenyi István University, in the Department of Logistics and Forwarding. Her research focuses on the structured design of logistics strategies, including the internal coherence of logistics strategies. Furthermore, she investigates the emergence of cognitive biases in decision-making processes in the logistics field.

APPENDIX 1



Overlay Visualization by VOSViewer from the research results of ScienceDirect on the 10th of April 2024. Keywords: "push pull" AND (logistics OR "supply chain management") AND sustainability. The minimum number of occurrences of a keyword: 2.

APPENDIX 2



Overlay Visualization by VOSViewer from the research results of Scopus on the 10th of April 2024. Keywords: "push pull" AND (logistics OR "supply chain management") AND sustainability. The minimum number of occurrences of a keyword: 20.

APPENDIX 3

Types of cognitive biases	Pattern recognition biases	Storytelling trap	Instinctively searching for information that provides justification to support a situation.
		Confirmation bias	Searching for corroborating information about pre-existing views, e.g. policy. Ignoring information that contradicts our point of view.
		Champion bias	The company trusts a well-performing professional more, even if the information is objectively not relevant.
		Experience bias	To make a decision based on memory and experience in an environment where these experiences are no longer applicable.
		Attribution error	Success (or failure) is attributed to a single person in the company. E.g. Steve Jobs, Apple
		Halo effect	Adopting the best practices of a successful person in the hope that the strategy they use will work.
		Survivorship bias	Only the strategies of successful leaders are studied. There are several lessons to be learned from studying failed companies.
	Action-oriented biases	Overconfidence	To overestimate one's abilities, which affects the decisions made.
		Planning fallacy	Excessively optimistic planning, regarding budget and time.
		Overprecision	To overestimate our ability to predict the future.
		Competitor neglect	Underestimating competitors, ignoring the competition
	Cognitive inertia	Anchoring	The leader tends to use the numbers presented to him as an "anchor" even if that number is not relevant to the case. For example, an annual budget plan.
		Commitment escalation	Because of a promise, they won't change the already established strategy so that the energy invested until then "doesn't go to waste"
		Status quo bias	Avoiding difficult decisions, therefore maintaining the status quo.
		Loss aversion	Loss aversion is one of the biggest barriers to development.
		Uncertainly aversion	Avoiding unknown risk in order to avoid loss.
		Hindsight bias	They judge the same occurrence with different probability before and after the event has occurred.
	Self-Interest biases	Present bias	The company's management does not think long-term. Immediate benefits instead of future profits.
		Self-serving bias	Individuals are driven by an unconscious intention to make decisions for their own self-interest, whether financial or emotional.
		Omission bias	The management of the company accepts when someone else makes a mistake and not themselves. In addition, they consider it morally acceptable to profit from it.
	Group biases	Groupthink	People tend to adopt the collective point of view even when they know it is not correct.
		Information cascades	The order of speakers distorts the outcome, as the opinions of those who speak first are amplified.
		Polarization	As a result of groupthink, the group's opinion will be more extreme, which the group members will represent more confidently. It also deepens commitment.

Design Suggestions for Digital Workflow Oriented Desktop VR Spaces

Anna Sudár[1,2,3,*], and Ildikó Horváth[1,2,3,*]

Abstract—**This paper presents an examination of design principles for 3D Virtual Reality (VR) environments, with a focus on enhancing digital workflows. Employing objective data, the study sets out to clarify the primary design considerations for crafting effective 3D VR spaces. Through empirical research, the authors conducted comparative analyses of task performance within both classical 2D Windows and in 3D VR environmental contexts, exploring users' perceived difficulty levels alongside eye-tracking data. The findings reveal that, although 3D VR environments rich in distracting elements and demanding high navigational effort increase perceived task difficulty, these factors do not negatively impact overall performance or task completion time. Interestingly, eye-fixation duration results indicate that visual fixation in 3D VR falls within expected norms, whereas in 2D scenarios, fixation rates are significantly higher, more than doubling those observed in 3D settings. Drawing on these insights, the paper supports the design of 3D VR spaces that are simpler and intuitive, necessitating minimal navigation, thereby optimizing task performance efficiency.**

Index Terms—**virtual reality, desktop VR, virtual space design**

I. Introduction

Ergonomic considerations in Virtual Reality (VR) are crucial for ensuring a comfortable, safe, and engaging user experience and supporting performance efficiency. These considerations primarily address the physical interaction between the user and the VR environment to minimize discomfort and potential health risks. A large volume of scientific studies support the idea that user experience (UX) design is crucial, as good UX design has been shown to enhance engagement and motivation, and can help maintain user attention for longer durations compared to traditional 2D interfaces [1]–[3]. These effects also extend to 3D spaces, as they allow users to create, visualize, and recall information in visually appealing and persuasive learning environments [4], [5]. Desktop VR combines elements of traditional computer use with the immersive qualities of VR, making it essential to consider both digital ergonomics and the unique demands of a 3D virtual environment. In today's digital interface design, which is connected to task performance and knowledge acquisition, it is advisable to take into account research results based on eye-tracking technology measurements. These studies have found correlations between fixation numbers and task difficulty, and fixation durations might be influenced by underlying affective processes that contribute to learning. [6]–[9].

[1] Corvinus Institute for Advanced Studies, Corvinus University of Budapest, Hungary
[2] Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Hungary
[3] Hungarian Research Network, HUN-REN, Budapest, Hungary
(*Email: {anna.sudar, ildiko.horvath}@uni-corvinus.hu)

DOI: 10.36244/ICJ.2024.5.14

The UI should be intuitive and easy to navigate. One primary consideration in designing spatial elements is to take into account human spatial perceptions. Accordingly, it is recommended to design virtual spaces as open areas or, in the case of closed spaces, with high ceilings [10]. Good quality audio that is synchronized with the visual elements of the VR environment can enhance the experience without causing auditory strain. Text and icons need to be large and clear enough to be easily readable. Interaction methods should be natural [11] and not require excessive or uncomfortable movements or complicated workflows. The ability for users to customize their VR experience (like adjusting sensitivity settings, UI elements, etc.) can greatly improve comfort and accessibility for a wide range of users. In this study, drawing from our research results, we aim to establish design guidelines aimed at enhancing user productivity and working effectiveness and contributing to the advancement of desktop VR environments. The structure of the paper is organized as follows: In the second section, the Definition and Metrics are presented. This section introduces key concepts, definitions, and the metrics used to evaluate virtual reality environments. The next section is the intersection of Cognitive Infocommunications (CogInfoCom) [12], [13] and cognitive aspects of VR (cVR) [14] technologies, presenting relevant research findings. The Research Context sets the background for understanding the framework of this study. Then, the authors present the main finding of the study including data analysis, interpretations, and the implications of research within the VR domain. The session Discussion critically examines the research results, discussing their significance, the limitations encountered, and how they relate to existing literature. Based on this research, the authors propose actionable recommendations for designing immersive and user-friendly VR environments to support performance efficiency. The paper concludes by summarizing the key findings, their implications for the field of VR, and suggesting directions for future research.

II. Definition and metrics

The objective of this chapter is to define essential terminology, providing clarity and ensuring a comprehensive understanding of the terms employed throughout the paper. Several of the key definitions presented herein have their origins in prior publications by the authors. However, it is imperative to emphasize that their application within this context is entirely novel and original.

A. Digital workflow—DW

[15] Digital workflows determine the order in which individual digital elements are to be accessed or processed during the course of a digital project. We distinguish among the following types of digital workflows:

1) 1st order (linear): The digital elements are to be accessed in a static and sequential order, one after the other.

2) 2nd order (loopy): There are loops in the order in which the digital elements are to be accessed, so that individual elements, or smaller sequences thereof, are to be accessed repetitively. Such loops can be characterized by length and number of repetitions.

3) 3rd order (networked): Digital elements accessed during the project are structured as hierarchical loops, so that the project may contain subprojects of subprojects, and/or the ordering of digital elements may be different upon different repetitions of the loops.

4) 4th order (algorithmic): It is possible that the project contains branches, so that different digital elements are accessed dynamically in an order that depends on information obtained during the project.

B. Digital Guidance—DG

[15] Digital guidance is taken to mean a process that unambiguously drives the user's attention during the digital workflow and thus reduces (partially, or to 0) the time required for searching for and finding the relevant digital content. It is possible to distinguish among three forms of digital guidance as follows:

1) none: no guidance is applicable, or the representation of the digital content doesn't involve embedded digital elements (instead, the elements are provided through separate lists).

2) sequential (DG-S): The digital elements are traversed in sequential order. It is thus possible to jump from one element to the next in the context of a digital workflow.

3) random access (DG-R - event/dynamic focus-driven): One can switch between sequences of digital elements, and thus follow non-static sequences (for example, in the case of DWs of the 4th order).

C. Information Availability—IA

[11] This indicates what percentage of the information (digital content) needed to execute a workflow is available in the digital work environment when executing up to 1 Navigation Based elementary Operation.

Remark: e.g., 100 percent, if all the information required to execute the workflow is available and can be accessed by a navigation operation.

D. Information Access Cost—IAC

[11] The weighted sum of the time spent accessing information for each type of operation, where the weights are the number of elementary operations corresponding to that type of operation.

$$IAC = \sum_{i=1}^{n} O_i * t_i \quad (1)$$

Remark: this metric is high even when complex operations are performed in the same amount of time and when simple operations are performed for a long time. This includes the complexity of the operation and the user's ability.

E. Information Validity—IV

[11] This indicates the percentage of the information presented in the digital work environment that is directly required for task execution.

IV = Number of Valid Information Units/Summa necessary information units number.

F. Personalized Workflow Order Ability—PWO

[11] This is an indicator of the facility of the digital work environment to provide users with the ability to arrange digital content in their own way. The value is 0 if the option is not provided, 1 if the users can set the layout themselves, and 2 if the optimal layout is automated with AI support.

G. Personalized Information Overview—PIO

Indicates the ability of the digital work environment to provide users with the ability to set personalized information overview. The value is 0 if the option is not given, 1 if the users can set the layout themselves, and 2 if it is automated with AI support.

H. Preference point

Preference points in the virtual space that users were more likely to visit in order to find the spots that best allowed them to oversee the space and solve their tasks.

I. Content arrangement types

Patterns with which users preferred to arrange the content

1) Content: content types that had a similar subject matter were most likely to be arranged in clusters, in close proximity to each other.

2) Type: similar content types were most likely to be arranged in clusters, close to each other.

3) Mixed: primal organizing principle is the content, but within the same area, the secondary organizing principle is the type.

III. Antecedent researches by CogInfoCom and cVR

CogInfoCom stands at the intersection of infocommunications and cognitive sciences, striving to enhance the co-evolution and interplay between human cognition and digital technology. This interdisciplinary field is dedicated to elevating human efficiency in digital workspaces and refining work processes through innovative IT solutions. It delves into the mutual evolution of infocommunication devices and human cognitive functions, aiming to optimize interaction within digital environments [12], [13], [16].

cVR, on the other hand, delves into the expansion of human cognitive abilities through the utilization of various technologies within a three-dimensional spatial framework.

Here, VR systems, as advanced infocommunication devices, play a pivotal role. They revolutionize information organization and management, allowing for spatial categorization and prioritization of data according to its significance in the workflow. VR's capacity for collaborative use and integration of other technological advancements further underscores its value in this domain [14].

CogInfoCom and the cVR field are extensively researched scientific domains, showcasing significant advancements such as the development and enhancement of the three-dimensional virtual library model [17], [18]. This model opens up new possibilities in digital information architecture. The fields also delve into 'mathability,' which investigates the synergy of artificial and natural cognitive abilities in mathematics [19], [20]. Furthermore, they are at the forefront of pioneering new educational methodologies, or foreign language education and linguistics [21], and are actively exploring learning challenges and emerging opportunities in the rehabilitation of autistic children [22], underlining their expansive and dynamic nature.

Moving to the connection with UI design, VR systems herald a paradigm shift from conventional command-based interactions to dynamic, user-centric interfaces. This shift necessitates a deep understanding of human behaviour within these virtual environments, provided by cognitive science. The inherent properties of VR, coupled with insights into human cognition, pave the way for more sophisticated and efficient virtual workspaces compared to traditional 2D interfaces. The introduction of a third dimension not only facilitates more organized information management but also resonates with the innate spatial understanding of human users. This spatial familiarity, along with the use of metaphors, contributes to a robust comprehension of tasks and data. Consequently, desktop VR systems enable collective visualization and observation, fostering an environment conducive to the sharing of knowledge and information [23].

VR brings new challenges and opportunities in the context of UI and information management. While traditional Windows interfaces offer users a familiar, structured environment controlled by mouse and keyboard, VR is a digital interface that provides interaction in a much more immersive spatial environment and operates according to a completely different paradigm. In recent years, several publications in the field of CogInfoCom and cVR have presented research results related to VR design. Of these, the relevant results for this study are:

Virtual Reality (VR) introduces a dynamic array of challenges and possibilities within the realms of User Interface (UI) design and information management. Contrasting the conventional, well-known setup of Windows interfaces, which users navigate via keyboard and mouse, VR offers a novel and immersive spatial interface operating under a fundamentally different paradigm. Recent years have seen a surge in scholarly work within Cognitive Infocommunications (CogInfoCom) and Cognitive Aspects of Virtual Reality (cVR), contributing significant insights into VR design.

Experimental research has shown that preference points and attentional focus points can be found in 3D virtual reality spaces. In three-dimensional virtual reality settings, users tend to concentrate more and spend additional time at certain points of interest, which helps them navigate and complete assigned tasks [24]. Moreover, using virtual reality interfaces has been linked to improved recall of the process involved in organizing three-dimensional objects in a given space [25], [26], and also enhances the retention of information distributed across a space [25]. Cannavò and colleagues suggest that automating the conversion from two-dimensional to three-dimensional formats can significantly enrich the user experience in virtual reality workspaces, making it more engaging and productive [27].

Berki and associates conducted a study comparing how well users remember images shown in a three-dimensional virtual reality setting as opposed to a two-dimensional website. They discovered that the virtual reality environment was more effective in aiding memory recall [28], [29]. In a similar vein, it was observed that desktop virtual reality systems outperform traditional two-dimensional browsers in memory retention of additional information [30].

In previous research conducted by one of the contributors to this study, the effectiveness of personalization based on learning styles in three-dimensional desktop virtual reality was examined. The findings from that research revealed that when the instructional content was tailored to their individual learning styles, users achieved scores that were 20 percent higher, along with an 8–10 percent improvement in response times in subsequent assessments [31]. This customization of the learning environment demonstrated a significant impact on user behaviour and performance in the 3D virtual reality context.

These studies collectively indicate that virtual reality technology may play a crucial role in reducing the mental effort required for spatial memory and in improving performance in tasks that involve spatial orientation, particularly those that involve navigating through a vast array of digital documents.

In virtual spaces, the creation of a sense of presence is of utmost importance, and it is equally crucial to bestow upon the user a feeling of control [32]–[34]. This serves several purposes. Firstly, in reality, individuals have control over their own bodies, primarily altering their positions, orientations, and perceptions of their environment [35], [36]. This agency must be replicated within virtual worlds to ensure the user's sense of security. Additionally, research has shown that active participation and control over one's environment lead to better retention of events and information [37], [38]. Thus, providing users with the ability to actively shape their virtual surroundings enhances their sense of control.

The concept of control is closely tied to perception, which, in the real world, flows through various modalities such as visual, auditory, and haptic channels. These modalities aid in navigation, information processing, storage, and even survival [39]. In the realm of virtual reality, almost all these modalities can be simulated. However, haptic feedback remains less developed in desktop VR, lagging behind the tactile sensations provided by controllers, gloves, and similar devices. Additionally, replicating olfaction, the sense of smell is currently a limitation in virtual technology. Nevertheless, groundwork has been laid to address these challenges and further enhance sensory immersion in virtual environments [40]–[42].

Special Issue
of the Infocommunication Journal

IV. Research Background

A. Subjects

A total of 21 participants were divided into two groups for each condition, and no specific qualifications or expectations were considered when selecting participants. For the 3D VR measurement, 14 participants initially took part, but due to technical issues and breaks in data collection, the results from nine subjects (three women and six men, aged between 17–55 years, with a mean age of 32.5 years) will be published in the analysis. The 2D environment study involved seven participants (four women and three men, aged between 25–33 years, with a mean age of 27.83 years). All participants were Hungarian native speakers, volunteering for the experiment with informed consent obtained in advance. The research was conducted under institutional endorsement ensuring ethical compliance and data privacy.

B. Procedure

This study explores some of the cognitive capabilities of users in desktop VR environments by comparing spatial behaviours, performance, and subjective experiences between traditional 2D interfaces and immersive 3D VR settings. At the initiation of the session, participants engaged in the prescribed task within three dimensions were requested to demonstrate their familiarity with the MaxWhere software and provide an estimation of the time they had allocated to using the application. Those participants who lacked prior knowledge of the software practised for around 30 minutes in acquainting themselves with its functionalities and acquiring fundamental user capabilities. Proficiency in essential skills encompassed adept navigation within the software, as well as the ability to activate and deactivate the display panels and engage with the content presented on them. The evaluators assessed the mastery of proficient software utilization. Participants were tasked with studying materials related to astronomy and completing questionnaires in either a 2D environment hosted on Google Sites or a 3D environment using the MaxWhere desktop VR platform. Despite using identical study materials and questionnaires in both settings, they were presented differently: linearly in the 2D interface and spatially in the 3D VR environment as shown in Figure 1. The questionnaires that were specifically designed to align with each of the subtopics were composed of a series of true-or-false questions, a set of multiple-choice questions, as well as a collection of questions that necessitated brief responses typically consisting of only one or two words. The task was a reading comprehension exercise which is very common in education and also in work scenarios, where participants found all the necessary information in the digital materials placed around the questionnaires. These materials were essential for answering the questions posed in the questionnaires. One example of the true-or-false question type: "Black holes can be observed through their gravitational effects on the surrounding gas, dust, and stars. True or False?" - The participants were asked to find the information around the questionnaire in on of the related digital contents and mark in the questionnaire the right answer.

The theme of all questionnaires revolved around astronomy. The four subthemes of the questionnaires were ("*Universe*", "*Planets*", "*Satellites*" and "*Space Research*"). We chose this topic because none of the participants were experts in the field, ensuring that they approached the tasks with similar levels of knowledge. This prevented any significant disparities in results caused by someone being highly knowledgeable in certain areas and potentially skewing the measurement outcomes.

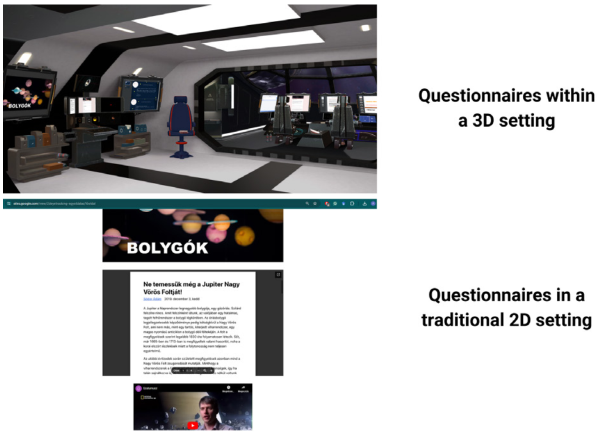


Fig. 1. This Figure illustrates through an example that the same digital content and questionnaires appeared for participants in both the 2D and 3D scenarios. The difference was visibly apparent in the layout mode.

C. Technical background

We used an HP Omen 15 laptop for conducting the research, which had the following specifications:

- AMD Ryzen™ 7 5800H processor
- NVIDIA® GeForce RTX™ 3060 graphics card
- 15.6-inch screen
- 16 GB RAM
- 512 GB SSD

The device used for Eye Tracking measurements was an EyeTribe eye tracker along with OGAMA (Open Gaze and Mouse Analyzer) Version 5.1 software. Since the VR software used for the measurement was a desktop VR application, it did not require the use of extra peripherals (such as an HMD). The eye-tracker was positioned at the bottom of the screen at an appropriate angle, ensuring accurate and consistent data collection for each individual.

D. Experimental environment design

When designing the virtual environment, special attention was paid to incorporating the experiences gained from previous research into the 3D virtual space created for this measurement. Ensuring dynamic interaction with the content was an important consideration because we assumed that such a setup would better meet the demand for a holistic overview of the educational materials, similar to placing documents on a physical desk. This could potentially lead to more effective learning outcomes and more effective performance compared to the more static and linear presentation of materials in the 2D environment.

The experiment was designed with several key objectives structured around both quantitative and qualitative assessments:

Objective Performance Metrics, This includes quantifiable data such as:

- The total number of correct answers participants provided.
- The amount of time participants spent completing the questionnaires.

Subjective Assessment: This assessment was based on the participants' personal evaluations of:

- The difficulty level of the questionnaires.
- Interaction Patterns in a 3D Environment: The study also examined how participants interacted within a three-dimensional (3D) environment, focusing on specific behaviours and engagement methods.

V. Research Results

A Mann-Whitney U test was used to compare the scores of the 3D Group (Mdn = 614.07) and the 2D Group (Mdn = 317.16) (descriptive statistic in Fig. 1) on the fixation duration mean. The Mann-Whitney U statistic was U = 55, indicating a significant difference between the two groups ($p = 0.012$) (Fig. 2).

	Fixation Duration Mean (ms)	
	2D	3D
Valid	7	9
Missing	0	0
Mean	766.390	311.948
Std. Deviation	558.407	42.459
Minimum	289.074	254.968
Maximum	1916.796	392.029

Fig. 2. Descriptive statistics of the fixation duration mean between the two measured groups.

We conducted an independent samples t-test to assess the differences between 2D and 3D environments across various thematic sections. Each analysis was preceded by a Levene Test for Equality of Variances, confirming homogeneity in variances across comparisons. The results indicated no significant differences between the groups, except in the time required to complete the satellite-themed questionnaire. Participants in the 3D environment completed this task noticeably quicker (units of measure were minutes), showing a significant difference ($t(14) = 3.38$, $p < 0.05$).

In the 3D virtual space, participants found the sections on Satellites and the Universe the most challenging, with an average score of 4.25 out of 7 (SD: 1.6). This was followed by the Planets section with a score of 3.75 (SD: 1.42), and

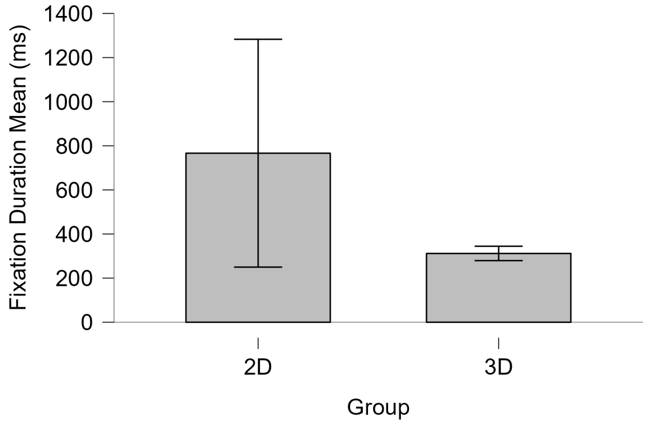


Fig. 3. Fixation duration mean differences between the 3D VR and the 2D group.

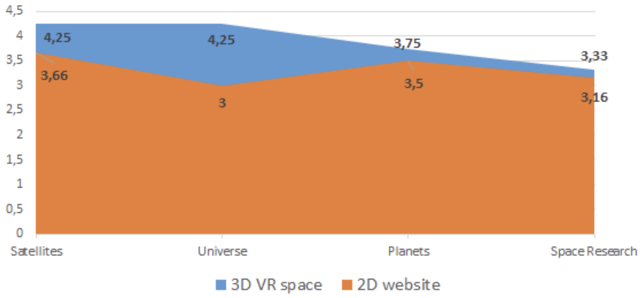


Fig. 4. Subjective user feedbacks - Comparative difficulty and performance in 2D vs. 3D virtual learning environment.

the Space Research section was the easiest at 3.33 points (SD: 2.01). Conversely, participants navigating the 2D website rated the Satellites section as the most difficult, scoring an average of 3.66 (SD: 1.63), followed by Planets at 3.5 (SD: 1.22), Space Research at 3.16 (SD: 1.94), and the Universe section as the easiest with a mean score of 3 (SD: 1.095). Fig 3. shows the research results connecting to these user feedbacks.

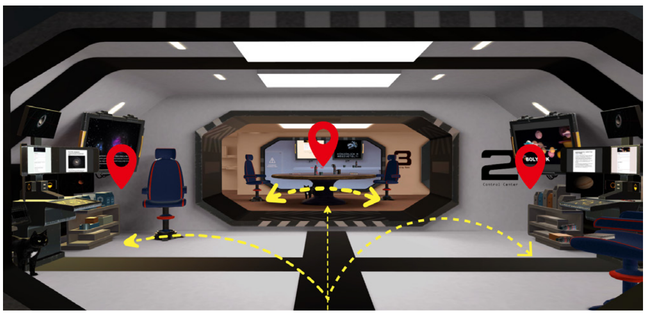


Fig. 5. Interaction routes and overview point examples.

Based on the analysis of screen recordings, a prevalent interaction patterns (Figure 5.) observed among the participants in the 3D measurement was what we refer to as the "overview mode". The red marks in Figure 5. indicate some examples

Recommendation	Description	Relevant research results
Improve Multimodal Stimulus and Feedback	Use auditory and visual contents and feedback to enhancing immersion and information delivery.	- Subjective assessment of the difficulty level of the questionnaires - 2D 3D differences and content analysis
Ensure Consistency	Apply consistent UI/UX design principles, accommodating dynamic VR interactions and perspectives.	- Information overview - Fixation duration time results
Minimize Textual Instructions	Prioritize visual and spatial cues over textual content for guidance, supporting seamless user engagement.	- Subjective assessment of the difficulty level of the questionnaires - Fixation duration time
Incorporate Familiar Elements	Utilize familiar interfaces and interaction models from 2D environments to improve VR usability and ease the transition.	- Overview interaction - Fixation duration time.
Strategic Content Placement	Organize digital content for easy navigation and task management, using circular arrangements and prominent viewpoints.	- Overview, alternating view.
Optimize Object and Content Placement	Consider the size, type, and spatial relationships of content and objects to support tasks without excessive cognitive load.	- Fixation duration time - Overview and alternating view

Fig. 6. Design recommendations for workflow-oriented desktop VR spaces.

of the overview points and the yellow paths represent the interaction pathways throughout the space, that we identified based on the video analyses. In this mode, users sought a position for each block where all materials related to the block were well-visible and comprehensible. Additionally, an alternating pattern was also evident, similar to displaying parallel windows in 2D. In this case, users held two or more windows in a single view and toggled between active windows with clicks to perform tasks, thus saving time, navigation efforts, and energy.

VI. Discussion

The longer one needs to focus on a particular area, the greater the mental effort required for the user to complete the task or other common activity. In the presented measurement, the mean fixation duration of individuals participating in 2D and 3D measurements was observed. Upon analysis, we found that the mean fixation duration of individuals solving tasks in the 3D VR space was significantly lower than the control group working on the traditional 2D homepage. At the end of each questionnaire, participants were asked to rate the difficulty of the task block on a Likert scale, and at the end of the measurement, a closing questionnaire was administered asking which questionnaire they found most challenging. The questionnaires revealed that although individuals solving VR tasks rated each block slightly more difficult than the control group solving tasks in 2D, there was no significant difference in correctness of the answers and in completion time between the two groups. Furthermore, the mean fixation duration did not support this subjective evaluation; in fact, it indicated the opposite. At the beginning of the measurement, all users were familiar with the use and basic functioning of the 3D software. However, it became clear during the survey and evaluation of the results that due to the layout and design of the space, there was a need for precise navigation between individual blocks in the virtual space, and they were not able to find only a few preference points to see through the whole space, which could negatively affect the user experience. Additionally, observing the users' movements in the space, it is evident that the size and placement of the displays holding digital content in the 3D space also require reconsideration in the future. Although the questionnaires were always centrally located with the necessary content around them, the size and placement varied within the blocks. It would be advisable in the future to restructure these elements so that users are assisted by similar layouts during task completion. For example, videos, PDFs, etc., could be located in the same position in each block and of similar size.

Furthermore, a recurring pattern was that users treated the images containing the titles of blocks as content rather than labels. Therefore, another suggestion for future design is to fully separate the titles/content elements of the content blocks to prevent users from attempting interaction with them and to serve solely as information. Interpreting the results and analyzing user behavior suggests that further optimization and modification of the designed 3D virtual space are necessary to maximize the user experience. To achieve this, we propose the following guidelines.

VII. Designing the Future: Guidelines for VR Development

In virtual environments, user experience hinges on control and presence, crucial elements highlighted by research [32]–[34]. Users' ability to exercise agency in virtual spaces mirrors their actions in the real world, fostering engagement and a sense of security [35], [36]. Active participation and control not only enhance user engagement but also significantly contribute to improved information retention [37], [38]. While tactile feedback in virtual reality may not fully replicate real-world sensations, strategic implementations such as haptic feedback through controllers or gloves aim to enrich the immersive experience [43]. Similarly, auditory cues complement visual elements, aiding users in understanding actions and enhancing their sense of presence [44], [45].

Consistency in design is paramount across various viewing angles and orientations in virtual environments to ensure a seamless user interaction [46], [47].

In previous research within the field of CogInfoCom an cVR, studies [24], [48] have identified key preference points in desktop VR workspaces. In addition, the current study's finding of a significantly shorter fixation duration time in 3D VR underlines the importance of spatial instructions and visual signals in navigating virtual environments. This suggests immersive experiences are more effective than extensive textual guidance [49], [50] in enhancing user commitment. Furthermore, incorporating familiar design elements into these spaces not only increases user comfort but also facilitates task completion [51], [52] [50], [51].

Based on the results of this study on interaction patterns and previous research results of the author regarding the workflow oriented VR spaces, we conclude that effective grouping and clustering of content are essential in virtual workspaces to enhance user experience and effectiveness [24], [48]. Circular arrangements of content groups support holistic overview modes, aiding task monitoring and navigation [24], [53]. Considering the size of virtual spaces is vital to prevent cognitive overload and time loss, with predefined spatial elements facilitating user preferences and customization [53].

Content types, including PDFs, images, presentations, videos, and web content, serve varying roles in virtual environments. Users tend to display PDFs on monitors and videos on projector screens, with display orientation impacting user preference, favoring vertical or slightly inclined displays [53]. Introducing flexible display panels allows users to customize layout, number, size, and relationships of placed content, further enhancing user interaction and customization [53].

Summarizing the main findings above, Figure 6. shows the design recommendations of workflow-oriented VR spaces based on the research results of the current study.

By adhering to these recommendations, VR designers can create spaces that are not only immersive and engaging but also intuitive and efficient for users, leveraging the unique capabilities of VR while addressing its current limitations.

VIII. Conclusion

In conclusion, this study highlights the importance of finding the right balance between providing relevant information and managing potential distractions in 3D VR environments. While the inclusion of detailed 3D animations and well-organized content blocks can enhance immersion, it also introduces complexities that may hinder user experience. Despite users' feelings of increased difficulty, objective eye-tracking data suggest that users navigate 3D VR environments effectively, even in the presence of distractions.

Moving forward, further research is needed to refine strategies for optimizing information presentation in 3D VR environments, with a focus on improving user experience and task performance. Simplifying design principles and prioritizing intuitive navigation can help maximize efficiency and user satisfaction in future 3D VR environments. This study underscores the importance of ongoing research and iterative design processes to fully realize the potential of 3D VR technology in the field of the design of digital workflow-oriented desktop VR spaces.

Acknowledgments

The research presented in this paper was supported by the Hungarian Research Network (HUN-REN), and was carried out within the HUN-REN Cognitive Mapping of Decision Support Systems research group.

References

[1] A. Afrooz, L. Ding, and C. Pettit, "An immersive 3d virtual environment to support collaborative learning and teaching," *Computational Urban Planning and Management for Smart Cities 16*, pp. 267–282, 2019. DOI: 10.1007/978-3-030-19424-6_15

[2] N. Pellas, S. Mystakidis, and A. Christopoulos, "A systematic literature review on the user experience design for game-based interventions via 3d virtual worlds in k-12 education," *Multimodal Technologies and Interaction*, vol. 5, no. 6, p. 28, 2021. DOI: 10.3390/mti5060028

[3] J. M. Evangelista Belo, A. M. Feit, T. Feuchtner, and K. Grønbæk, "Xrgonomics: facilitating the creation of ergonomic 3d interfaces," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11. DOI: 10.1145/3411764.3445349

[4] Á. B. Csapó, I. Horvath, P. Galambos, and P. Baranyi,"Vr as a medium of communication: from memory palaces to comprehensive memory management," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 2018, pp. 000 389–000 394. DOI: 10.1109/CogInfoCom.2018.8639896

[5] B. Berki, "2d advertising in 3d virtual spaces," *Acta Polytechnica Hungarica*, vol. 15, no. 3, pp. 175–190, 2018. DOI: 10.12700/APH.15.3.2018.3.10

[6] S. Negi and R. Mitra, "Fixation duration and the learning process: An eye tracking study with subtitled videos," *Journal of Eye Movement Research*, vol. 13, no. 6, 2020. DOI: 10.16910/jemr.13.6.1

[7] A. Poole and L. J. Ball, "Eye tracking in hci and usability research," in *Encyclopedia of human computer interaction*. IGI global, 2006, pp. 211–219. DOI: 10.4018/978-1-59140-562-7.ch034

[8] R. N. Meghanathan, C. van Leeuwen, and A. R. Nikolaev, "Fixation duration surpasses pupil size as a measure of memory load in free viewing," *Frontiers in human neuroscience*, vol. 8, p. 1063, 2015. DOI: 10.3389/fnhum.2014.01063

[9] B. Albert and T. Tullis, *Measuring the User Experience: Collecting, Analyzing, and Presenting UX Metrics*. Morgan Kaufmann, 2022. ISBN 0124157920, 9780124157927

[10] S. H. Cha, C. Koo, T. W. Kim, and T. Hong, "Spatial perception of ceiling height and type variation in immersive virtual environments," *Building and Environment*, vol. 163, p. 106285, 2019. DOI: 10.1016/j.buildenv.2019.106285

[11] I. Horváth and B. Berki, "Investigating the operational complexity of digital workflows based on human cognitive aspects," *Electronics*, vol. 12, no. 3, 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/3/528 DOI: 10.3390/electronics12030528

[12] P. Baranyi and Á. Csapó, "Definition and synergies of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.

[13] P. Baranyi, A. Csapo, and G. Sallai, *Cognitive infocommunications (coginfocom)*. Springer, 2015. DOI: 10.1007/978-3-319-19608-4

[14] I. Horvath, Á. B. Csapó, B. Berki, A. Sudar, and P. Baranyi, "Definition, background and research perspectives behind 'cognitive aspects of virtual reality'(cvr)," *Infocommunications Journal: A Publication of the Scientific Association For Infocommunications (HTE)*, no. SP, pp. 9–14, 2023. DOI: 10.36244/ICJ.2023.SI-IODCR.2

[15] I. Horváth and A. Sudár, "Factors contributing to the enhanced performance of the maxwhere 3d vr platform in the distribution of digital information," *Acta Polytechnica Hungarica*, vol. 15, no. 3, pp. 149–173, 2018. DOI: 10.12700/APH.15.3.2018.3.9

[16] P. Baranyi, P. Galambos, Á. Csapó, and L. Jaloveczki, "Cognitive navigation and manipulation (coginav) method," Feb. 1 2018, uS Patent App. 15/658, 579.

[17] I. K. Boda, E. Tóth, M. Bényei, and I. Csont, "A three-dimensional virtual library model of the ancient library of alexandria," 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:56215123 DOI: 10.14794/ICAI.9.2014.1.103

[18] I. K. Boda, E. Tóth, I. Csont, and L. T. Nagy, "Toward a knowledge base of literary content focusing on the ancient library of alexandria in the three dimensional space," *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 251–258, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:25320574 DOI: 10.1109/CogInfoCom.2015.7390600

[19] K. Chmielewska and A. Gilányi, "Mathability and computer aided mathematical education," in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2015, pp. 473–477. DOI: 10.1109/CogInfoCom.2015.7390639

[20] K. Chmielewska, A. Gilányi, and A. Łukasiewicz, "Mathability and mathematical cognition," in *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2016, pp. 000245–000 250. DOI: 10.1109/CogInfoCom.2016.7804556

[21] J. B. Kóczy and L. I. Komlosi, "Revisiting literacy: Changing learning paradigms in digital culture," 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:204873640

[22] C. S. Lányi and Á. Tilinger, "Multimedia and virtual reality in the rehabilitation of autistic children," in *Computers Helping People with Special Needs*, K. Miesenberger, J. Klaus, W. L. Zagler, and D. Burger, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 22–28. DOI: 10.1007/978-3-540-27817-7_4

[23] J. Li, L. Khoo, and S. Tor, "Desktop virtual reality for maintenance training: an object oriented prototype system (v-realism)," *Computers in Industry*, vol. 52, no. 2, pp. 109–125, 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361503001039 DOI: 10.1016/S0166-3615(03)00103-9

[24] A. Sudar and A. Csapa, "Interaction patterns of spatial navigation and smartboard use in vr workspaces," in *Accentuated Innovations in Cognitive Info-Communication*. Springer, 2022, pp. 149–166. DOI: 10.1007/978-3-031-10956-0_7

[25] E. D. Ragan, A. Sowndararajan, R. Kopper, and D. A. Bowman, "The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments," *PRESENCE: Teleoperators and Virtual Environments*, vol. 19, pp. 527–543, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:17361861 DOI: 10.2312/EGVE/JVRC09/121-128

[26] B. Csapó, I. Horváth, P. Galambos, and P. Baranyi, "Vr as a medium of communication: from memory palaces to comprehensive memory management," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2018, pp. 000 389–000 394. DOI: 10.1109/CogInfoCom.2018.8639896

[27] A. Cannavò and F. Lamberti, "A virtual character posing system based on reconfigurable tangible user interfaces and immersive virtual reality," in *Smart Tools and Applications in Graphics*, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:88479188 DOI: 10.2312/STAG.20181297

[28] B. Berki, "Better memory performance for images in maxwhere 3d vr space than in website," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 2018, pp. 000 281–000 284. DOI: 10.1109/CogInfoCom.2018.8639956

[29] ——, "Role of presence, memory and spatial ability in a desktop virtual reality," in Accentuated Innovations in *Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 79–97. DOI: 10.1007/978-3-031-10956-0_4

[30] ——, "Desktop vr as a virtual workspace: a cognitive aspect," *Acta Polytechnica Hungarica*, vol. 16, no. 2, pp. 219–231, 2019. DOI: 10.12700/aph.16.2.2019.2.13

[31] I. Horváth, "An analysis of personalized learning opportunities in 3d vr," in *Frontiers of Computer Science*, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:237587652 DOI: 10.3389/fcomp.2021.673826

[32] S. Weech, S. Kenny, and M. Barnett-Cowan, "Presence and cybersickness in virtual reality are negatively related: a review," *Frontiers in psychology*, vol. 10, p. 158, 2019. DOI: 10.3389/fpsyg.2019.00158

[33] S. Grassini, K. Laumann, and M. Rasmussen Skogstad, "The use of virtual reality alone does not promote training performance (but sense of presence does)," *Frontiers in psychology*, vol. 11, p. 1743, 2020. DOI: 10.3389/fpsyg.2020.01743

[34] N. Cooper, F. Milella, C. Pinto, I. Cant, M. White, and G. Meyer, "The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment," *PloS one*, vol. 13, no. 2, p. e0191846, 2018. DOI: 10.1371/journal.pone.0191846

[35] A. Tapal, E. Oren, R. Dar, and B. Eitam, "The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment," *Frontiers in psychology*, vol. 8, p. 1552, 2017. DOI: 10.3389/fpsyg.2017.01552

[36] M. Velmans, "How could conscious experiences affect brains?" *Journal of Consciousness Studies*, vol. 9, no. 11, pp. 3–29, 2002.

[37] C. Pérez-Sabater, B. Montero-Fleta, M. Pérez-Sabater, B. Rising, and U. De Valencia, "Active learning to improve long-term knowledge retention," in *Proceedings of the xii simposio internacional de comunicación social*, 2011, pp. 75–79.

[38] J. G. Kooloos, E. M. Bergman, M. A. Scheffers, A. N. Schepens-Franke, and M. A. Vorstenbosch, "The effect of passive and active education methods applied in repetition activities on the retention of anatomical knowledge," *Anatomical sciences education*, vol. 13, no. 4, pp. 458–466, 2020. DOI: 10.1002/ase.1924

[39] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1385–1422, 1993. DOI: 10.1109/5.241504

[40] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger, "The impact of perception and presence on emotional reactions: a review of research in virtual reality," *Frontiers in psychology*, vol. 6, p. 26, 2015. DOI: 10.3389/fpsyg.2015.00026

[41] G. Freeman and D. Maloney, "Body, avatar, and me: The presentation and perception of self in social virtual reality," *Proceedings of the ACM on human-computer interaction*, vol. 4, no. CSCW3, pp. 1–27, 2021. DOI: 10.1145/3432938

[42] F. El Jamiy and R. Marsh, "Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality," *IET Image Processing*, vol. 13, no. 5, pp. 707–712, 2019. DOI: 10.1049/iet-ipr.2018.5920

[43] H. Benko, C. Holz, M. Sinclair, and E. Ofek, "Normaltouch and texturetouch: High-fidelity 3d haptic shape rendering on handheld virtual reality controllers," in *Proceedings of the 29th annual symposium on user interface software and technology*, 2016, pp. 717–728. DOI: 10.1145/2984511.2984526

[44] M. Hoppe, J. Karolus, F. Dietz, P. W. Woźniak, A. Schmidt, and T.- K. Machulla, "Vrsneaky: Increasing presence in vr through gait-aware auditory feedback," in *Proceedings of the 2019 CHI conference on human factors in computing systems,* 2019, pp. 1–9. DOI: 10.1145/3290605.3300776

[45] A. U. Batmaz and W. Stuerzlinger, "The effect of pitch in auditory error feedback for fitts' tasks in virtual reality training systems," in *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2021, pp. 85–94. DOI: 10.1109/VR50410.2021.00029

[46] S. Bodker, *Through the interface: A human activity approach to user interface design*. CRC Press, 2021. DOI: 10.7146/dpb.v16i224.7586

[47] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos, *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.

[48] A. Sudár and Á. Csapó, "Interaction patterns of spatial navigation in vr workspaces," in *2019 10th IEEE international conference on cognitive infocommunications (CogInfoCom)*. IEEE, 2019, pp. 615–618. DOI: 10.1109/CogInfoCom47531.2019.9089998

[49] T. Kojić, D. Ali, R. Greinacher, S. Möller, and J.-N. Voigt-Antons, "User experience of reading in virtual reality—finding values for text distance, size and contrast," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–6. **DOI**: 10.1109/QoMEX48832.2020.9123091

[50] P.-L. P. Rau, J. Zheng, Z. Guo, and J. Li, "Speed reading on virtual reality and augmented reality," *Computers & Education*, vol. 125, pp. 240–245, 2018. **DOI**: 10.1016/j.compedu.2018.06.016

[51] A. M. Cleary, A. S. Brown, B. D. Sawyer, J. S. Nomi, A. C. Ajoku, and A. J. Ryals, "Familiarity from the configuration of objects in 3-dimensional space and its relation to déjà vu: A virtual reality investigation," *Consciousness and cognition*, vol. 21, no. 2, pp. 969–975, 2012. **DOI**: 10.1016/j.concog.2011.12.010

[52] A. Almutawa and R. Ueoka, "The influence of spatial awareness on vr: investigating the influence of the familiarity and awareness of content of the real space to the vr," in *Proceedings of the 2019 3rd International Conference on Artificial Intelligence and Virtual Reality,* 2019, pp. 26–30. **DOI**: 10.1145/3348488.3348502

[53] A. Sudár and Á. B. Csapó, "Elicitation of content layout preferences in virtual 3d spaces based on a free layout creation task," *Electronics,* vol. 12, no. 9, p. 2078, 2023. **DOI**: 10.3390/electronics12092078

Ildikó Horváth obtained her PhD degree in 2017 and in 2023 completed her habilitation in Information Sciences. Her research focuses on the examination of the cognitive aspects of 3-dimensional alternative reality (VR, AR, Mixed reality) systems. The research goal is to determine the directions of development in digital ICT environments that ensure users' work with the most optimal cognitive load, thus positively impacting work efficiency. It also investigates the effects of digital environments on changes in human abilities and the possibilities of personalization. Shas experience in leadership, organization of education and development, and over 10 years of experience in scientific research, publishing, and university education for BSc, MSc, and PhD levels.

Anna Sudár obtained her PhD degree in 2023 in the field of computer science and holds an MSc degree in cognitive science. Her research topic is spatial cognition and the examination of general cognitive abilities in virtual reality. She has more than fifteen scientific publications including a book chapter. In the past 6 years as a scientific community activity, she participated in the organization of several international IEEE conferences and helped the editorial processes of many international Special Journal Issues.