Conceptualising immersive multimodal environments for psychomotor skills training.

Khaleel Asyraaf Mat Sanusi[a], Deniz Iren[b] and Roland Klemke[a,b]

[a] *Cologne Game Lab, TH Köln, Cologne, Germany*
[b] *Open Universiteit, Heerlen, The Netherlands*

Abstract

The coordination of psychomotor skills requires deliberate practice and techniques, all of which are typically taught in a physical setting, where instructions and timely feedback are given by the teachers. However, doing so remotely is commonly inefficient and ineffective, therefore, hindering the learner's progress. Sensors and immersive technologies enable the collection of multimodal data and the creation of immersion, respectively. These technologies have been widely used to further improve the learning outcome, especially in the psychomotor domain. In this paper, we present our research on designing an immersive training environment for remote psychomotor skill training and investigating how such an environment can be used for training skills in different psychomotor domains.

Keywords 1

Immersive technologies, Sensors, Multimodal, Psychomotor skills

1. Introduction

The global pandemic event of Covid-19 has affected various learning and teaching activities acutely. This necessitates the notion of online learning or e-learning in which web conferencing tools (e.g., Zoom, Teams) are widely utilised by teachers and students for classroom activities. However, this is rarely the case for psychomotor skills development as they require hands-on practice. Psychomotor skills need to be physically executed, in most cases, repetitively to the extent that the muscle memory is trained, which will automate the muscle movements [1]. Furthermore, the presence of teachers is needed in order to explain, demonstrate, and assess certain procedures. To achieve this, the human learning model has to be in a structured form where instructions are well-defined, and feedback can be given to ensure that the tasks are performed in a correct manner, which include safety and effectiveness. Timely and consistent feedback from the teacher is essential for the learner to avoid developing improper techniques during training, thus ensuring the desired goal can be achieved in a shorter time [2]. However, doing so in a remote manner makes the learning process ineffective and inefficient due to the lack of modalities such as haptic feedback or 3D full-body perception, hence impeding the learner’s progress. Due to this, psychomotor skill learners and teachers have been substantially affected.

Nowadays, educational technology and artificial intelligence (AI) researchers are progressively embedding sensor technologies for the collection of multimodal data, and machine learning approaches for tracking learners' behaviour and progress in authentic learning contexts. The combination of these technologies introduces new technological

Proceedings of the Doctoral Consortium of Sixteenth European Conference on Technology Enhanced Learning, September 20–21, 2021, Bolzano, Italy (online).
EMAIL: ks@colognegamelab.de (A. 1); deniz.iren@ou.nl (A. 2); rk@colognegamelab.com (A. 3)
ORCID: 0000-0001-6766-4416 (A. 1); 0000-0002-0727-3445 (A. 2); 0000-0002-9268-3229 (A. 3)

affordances that can be leveraged in the psychomotor education, especially in a remote manner, to further improve the learning outcome.

Multimodality is a theoretical assumption that can be applied to provide more structure in sensors for exploring learning. The general idea of multimodality in learning comes from the theory of embodied communication. Based on this theory, humans use their whole bodies to communicate with each other, applying various channels to exchange messages such as gestures, facial expressions, prosody, etc. [3]. Subsequently, the trend of multimodality has been employed in human-computer interaction. Sensor-based multimodal interfaces allow the monitoring of different modalities and have been applied in various domains to improve learning [[4], [5], [6]].

While the multimodal approach helps to improve the psychomotor learning outcome, designing virtual training environments adds immersion to the learning activity. As such, immersive learning technologies such as virtual reality (VR), augmented reality (AR) and game elements enable the creation of virtual training environments or simulations that typically consist of nearly, if not entirely, realistic physical similarity to an actual learning context. Herrington et al. [7] stress that the learning environment and designated tasks create the conditions for the “True” immersion. Hence, it can be argued that the instructions and feedback provided by the learning environment should be pragmatic for the learners to learn to perform the tasks in a correct manner; for example, personalised feedback (human-teacher-like) to create immersive learning experiences. Intrinsically, virtual training environments allow the learner to actively interact with the in-game objects which may create more engagement and increase motivation for the learner when performing tasks.

In this research, we aim to design and implement an immersive training environment for psychomotor skills using immersive technologies which will be integrated with sensor technologies and AI, in order to deliver instructions and feedback to learners in a meaningful manner. Furthermore, we intend to investigate the effectiveness of the system and whether it can be applied to train skills in different psychomotor domains. The development of this system provides an early and significant step towards combining immersive technologies and sensor technologies in a multi-sensor setup for collecting multimodal data and giving immediate feedback in an immersive training environment in which learners can use to improve their psychomotor skills independently.

The paper is structured as follows. In Section 2, we present related studies that utilise sensors for the collection of multimodal data and immersive training environments for immersion, and to what extent they are used in the psychomotor domain. Next, we explain our research questions in Section 3. Subsequently, in Section 4, we visualise and describe the research model and methods of this study. Finally, in Section 5 we discuss the expected outcomes of our study in theoretical and practical implications, followed by the conclusion.

2. Related work

2.1. Sensors in psychomotor learning

Sensor technologies are increasingly becoming more portable and increasingly used in psychomotor training, enabling efficient methods for the acquisition of performance data, which allows effective monitoring and intervention. That being said, such devices have been explored to provide support in the learning domain. For example, Schneider et al. [8] analysed 82 prototypes found in literature studies based on Bloom’s taxonomy of learning domains (psychomotor, cognitive, and affective). Their research suggests researchers and educators to consider utilising sensor-based platforms as reliable learning tools for reducing the workload of teachers and, therefore, contribute to the solution of many current educational challenges.

Motion sensors such as accelerometers and gyroscopes are predominantly used to acquire motion data to recognize human activities, especially in the psychomotor domain. These sensors are commonly combined and used in a synchronized manner to achieve a higher accuracy of detecting not only simple but complex activities as well [9]. This enables the collection of multimodal data and provides a more accurate representation of the learning process [10]. Furthermore, multimodal data can be collected using various sensors such as wearable sensors, depth camera sensors, Internet of Things devices, etc.

For instance, Schneider et al. [4] designed a system to support the development of public speaking skills using the Kinect v2 depth camera sensor to track the skeletal joints of the learner's body and the HoloLens headset to provide feedback in real-time when mistakes are detected while presenting. Limbu et al. [11] developed a system to teach basic calligraphy skills, which uses the pen sensor in Microsoft Surface tablet and EMG sensors in a Myo armband to provide feedback to learners during practice. It also allows the calligraphy teacher to create an expert model, which the learners can later use to practice and receive guidance and feedback based on the expert model.

To better understand learners' performance, educational researchers are progressively using machine learning approaches to classify activities based on the multimodal data collected. For instance, in the medical domain, Di Mitri et al. [5] investigated how multimodal data and Neural Networks can be used for learning Cardiopulmonary Resuscitation skills by utilising a multi-sensor system comprising of a Kinect v2 and a Myo armband. In the sports domain, Mat Sanusi et al. [6] applied the same framework as the previous author by using built-in accelerometer and gyroscope sensors in a smartphone and also a Kinect v2 to detect forehand table tennis strokes during training. Both study results show a high classification rate of the activities when combining the sensors, emphasising the importance of a multimodal approach in classifying complex activities.

In this research, we aim to use a multi-sensory system (e.g., wearable technologies, depth cameras) with the help of machine learning to help learners improve their psychomotor skills. We intend to have a theoretical framework that can be used for training skills in one psychomotor domain and subsequently applied in multiple domains.

2.2. Immersive training environments

Immersive learning technologies such as VR and AR are progressively becoming a significant medium for psychomotor training. Due to the substantial improvement and development in recent years, such technologies are being used in various psychomotor domains, including sports, physical training, rehabilitation therapy, and much more. These technologies transport individuals into an interactive training or learning environment, either virtually or physically, to replicate the authentic learning context of a specific skill.

For example, Song et al. [12] designed and implemented an immersive VR environment for teaching tennis using high-definition stereoscopic display, robust and accurate hybrid sensor tracking, shader-based skin deformation, intelligent animation control, and haptic feedback mechanism. The authors reported that, through these technologies, a real-time immersive tennis playing experience is achieved. Potentially, the system can be scaled to adapt various application cases such as other sports game simulations and even military training simulations.

Ali et al. [13] experimented with multiple VR fitness applications (e.g., VR Fitness, VirZOOM, BOXVR) for physical training such as walking, running, and jogging. In addition, they implemented a mobile application that uses built-in sensors such as an accelerometer and gyroscope for motion detection. As a result, they achieved up to 82.46% of accuracy and thus, described the effectiveness of VR technology in physical training, which is helpful for the development of psychomotor skills.

In our research, we aim to incorporate immersive technologies into the mix for the creation of immersive training environments to enhance the immersive experience of the learner in the learning setting. Our grand vision is to have a theoretical framework with a structured human learning model (feedback and instructions) within these immersive training environments that can be applied to not only one but also multiple psychomotor domains.

3. Research questions

Based on the problem identified and the related work analysed, we aim to investigate the following research questions (RQs):

1. What level of technological support (technology) is available in the literature and appropriate for delivering effective instructions and feedback

(pedagogy) to the learners in psychomotor training?

Fundamentally, it is crucial to identify *the most promising pedagogical approaches in psychomotor skills learning* that can be applied in multiple psychomotor domains. Furthermore, with technologies that have been widely used to improve the learning outcome in recent years, we survey *the state-of-the-art of technology that may potentially be helpful* for our research. Therefore, a systematic review will be carried out for these two processes and thus, answer our RQ1. The outcome of answering this question would be the theoretical framework of the system.

2. How can we create an immersive and information-rich (remote/self-learning) training environment for psychomotor skills that deliver effective instructions and meaningful feedback to the learner?

Subsequently, we design and implement a virtual training environment based on the theoretical framework retrieved from RQ1. The instruction and feedback systems should be given in a realistic manner to create immersive learning experiences. Therefore, it is vital to research *how can we maximise the system's effectiveness in providing feedback and instructions.* This includes the framing of interaction and the appropriate modalities for instructions and feedback. Consequently, we can investigate the effectiveness of the system: *can the system help learners improve their skills during training?*

3. To what extent can we generalise our training framework to multiple psychomotor domains?

Finally, we explore if the system can, both theoretically and practically, be adapted and applicable in multiple psychomotor domains. More exercise routines and common mistakes of the selected applications cases will be identified to suit the system's needs. Hence, it is crucial to know *what are other possible application cases that the new system can be used to train related psychomotor tasks* and *can the system effectively help learners learn different psychomotor skills?*

4. Methodology

4.1. Research methods

It is essential for this research to follow a methodological approach for designing, developing, testing, and evaluating such a system. Hence, we conduct our research based on the Design-based Research (DBR) approach, a common iterative methodological approach for prototypical solutions. In the context of our research, we combined two DBR models from Amiel & Reeves [14], and De Vielliers & Harpur [15], which are used in the domain of educational technologies.

Figure 1 shows the phases of the DBR approach for this research, and the following subsections explain each of the phases.

1. *Problem analysis:* In the first phase, a systematic literature will be reviewed to determine the importance of the problem and identify the current theory on the immersive multimodal environments in the psychomotor domain. Furthermore, the selection of application cases will be made in this phase. With these approaches, we are analysing the problem and defining research goals. The outcome of this step is a detailed research proposal containing goals and evaluation criteria.

2. *Design solution*: A theoretical framework is proposed based on the results from the systematic review, identifying the most promising pedagogical model in psychomotor training and the technologies that can be contributed to such a model. Our conceptual model (see Figure 2) states how we transfer the theoretical framework into our system design, suggesting to address the problem from phase 1.

3. *Develop solution*: The next phase is the implementation of the immersive training environment that serves the research purpose. The development of the system is based on the theoretical framework proposed in phase 2. The outcome is an innovative and functional immersive training environment system with the integration of immersive technologies, sensor technologies, and AI that aims to address the challenges of remote psychomotor training and help us achieve our research goals.

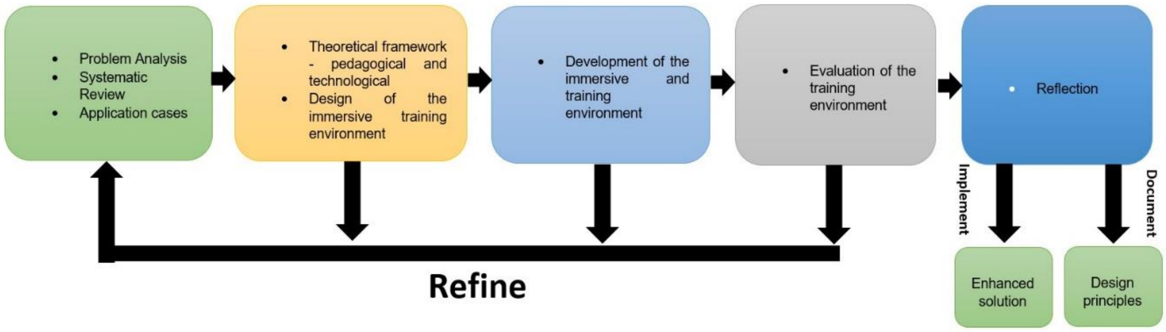

Figure 1: The synthesised model for DBR in the context of this research [[14], [15]].

4. ***Evaluate in practice***: Subsequently, in the next phase, focus group experiments involving the teachers/experts will be carried out for

qualitative analysis to gather important details that can be added to the system. Further, a user test will be conducted to reveal essential aspects of how the system can be improved. Additionally, questionnaires and surveys for the quantitative analysis are helpful to provide a general idea of how users perceive the interaction between the system. The refinement of the system should then be followed involving the teachers/experts to ensure that the system is ready to be tested with the learners in the real-world setting. Then, the data is collected and analysed to answer the research questions and to construct design principles.

5. ***Reflection, dual outcomes***:

Practical: This phase enhances the implementation of the solution. As reflection occurs, new designs can be further developed and implemented, which leads to an ongoing sub-cycle of the design-reflection process.

Theoretical: It is imperative to keep detailed records during the design research process concerning how the design outcomes (e.g., principles) have worked or have not worked, how the innovation has been improved, and what are changes have been made. Through this documentation, it can be helpful for other researchers and designers who are interested in those findings and examine them in relation to their context and needs.

4.2. Solution approach

In learning sciences, a conceptual model is commonly used to improve explanations and provide visual representations of abstracts [16]. Following this theory, we sketched a model for visualising the overall learning process using the immersive training environment from the human learner perspective (see Figure 2).

Based on the model, multimodal data will be collected by tracking the skeletal points and capturing the body motion of the human learner's body. Instructional tasks are ideally given before the learner performs the specific tasks. Feedback is typically given in real-time when mistakes are detected during training and as visual summative, after training. Instructions are also given during training to help learners progress to the next steps or even in the form of detailed feedback. These two aspects of the human learning model - instructions and feedback - can be given in multiple modalities. In the context of our research, the most common modalities that can be applied are visual, audio, and haptic. These modalities form various types of interaction that can be potentially used in the immersive training environment to give instructions and feedback such as virtual avatars, videos, etc. Finally, these aspects help validate the effectiveness of the immersive training environment.

Since the research is in an early phase, the conceptual model is still on the abstract level. However, this constitutes the groundwork of this research and will be extended into a bigger model with more aspects in the later phases.

5. Discussion and conclusion

The expected outcomes of this research are divided into two implications: theoretical and practical. From the theoretical perspective,

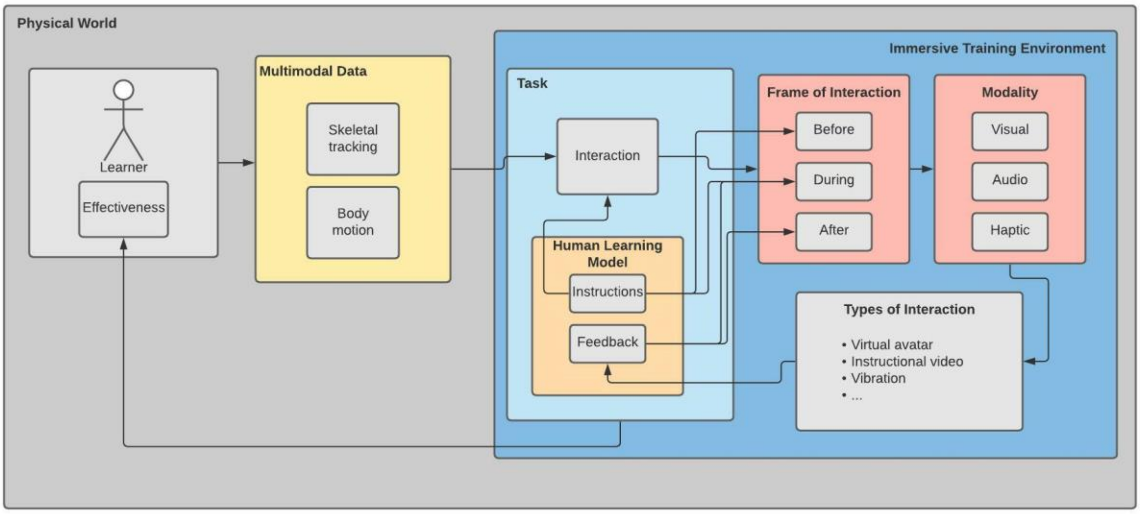

Figure 2: The conceptual model of this research.

systematic literature review findings on requirements to create immersive training environments for psychomotor skills will be delivered. Based on these findings, a conceptual framework of the immersive training environment consisting of guidelines and methodologies on delivering instructions, providing feedback, and tracking learner's performance will be constructed. We envision this framework to constitute the groundwork for the design. Moreover, it will extend immersive training environments for psychomotor skills training in multiple domains. This framework will potentially be useful for researchers as a basis for their theoretical and practical research.

From the practical perspective, a system for delivering effective instructions, providing meaningful feedback, and tracking learner's performance will be developed. Similarly, as the theoretical implication, such a system needs to be adapted in various psychomotor domains for different skills training. The empirical studies will be carried out with learners to measure the effectiveness of the system and the outcome should deliver promising results. Consequently, learners and teachers can benefit from the system to help them with the training.

This research investigates the effectiveness of an immersive training environment in the development of psychomotor skills training. The proposed theoretical framework integrates immersive technologies and sensor technologies for the immersion and multimodal data, respectively, providing a preliminary yet significant step towards combining such technologies in a multi-sensor setup to further improve the learning outcome in the psychomotor domain, especially in remote-learning scenarios.

6. Acknowledgements

This research is funded by the Federal Ministry of Education and Research (BMBF) in the program of Digital Higher Education for the project Multimodal Immersive Learning with Artificial Intelligence for Psychomotor Skills (MILKI-PSY) with the grant number: 16DHB4013.

7. References

[1] R. W. Barnes, Surgical handicraft: Teaching andlearning surgical skills, The American Journal ofSurgery 153 (1987) 422–427. Papers of the NorthPacific Surgical Association.

[2] A. Ericsson, M. Prietula, E. Cokely, The making of an expert harvard business review (july–august 2007), Expert Harvard Business Review,July–August (2007).

[3] I. Wachsmuth, M. Lenzen, G. Knoblich, Embodied communication in humans and machines, Oxford University Press, 2008.

[4] J. Schneider, D. Börner, P. van Rosmalen, M. Specht, Can you help me with my pitch? studying a tool for real-time automated feedback, IEEE Transactionson Learning Technologies 9 (2016) 318–327.

[5] D. Di Mitri, J. Schneider, K. Trebing, S. Sopka, M. Specht, H. Drachsler, Real-time multimodal feedback with the cpr tutor, in: I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, E. Millán (Eds.), Artificial Intelligence in Education, Springer International Publishing, Cham, 2020, pp. 141–152.

[6] K. A. Mat Sanusi, D. D. Mitri, B. Limbu, R. Klemke, Table tennis tutor: Forehand strokes classification based on multimodal data and neural networks, Sensors 21 (2021) 3121.

[7] J. Herrington, T. C. Reeves, R. Oliver, Immersive learning technologies: Realism and online authentic learning, Journal of computing in Higher Education 19 (2007) 80–99.

[8] J. Schneider, D. Börner, P. Van Rosmalen, M. Specht, Augmenting the senses: A review on sensor-based learning support, Sensors 15 (2015) 4097–4133.

[9] A. Dias Pereira dos Santos, K. Yacef, R. Martinez-Maldonado, Let’s dance: How to build a user model for dance students using wearable technology, in: Proceedings of the 25th Conference on User Mod-eling, Adaptation and Personalization, UMAP ’17,Association for Computing Machinery, New York, NY, USA, 2017, p. 183–191.

[10] P. Blikstein, M. Worsley, Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks, Journal of Learning Analytics 3 (2016) 220–238.

[11] B. H. Limbu, H. Jarodzka, R. Klemke, M. Specht, Can you ink while you blink? assessing mental effort in a sensor-based calligraphy trainer, Sensors19 (2019).

[12] P. Song, S. Xu, W. T. Fong, C. L. Chin, G. G. Chua, Z. Huang, An immersive vr system for sports education, IEICE TRANSACTIONS on Information and Systems 95 (2012) 1324–1331.

[13] S. F. Ali, S. Noor, S. A. Azmat, A. U. Noor, H. Siddiqui, Virtual reality as a physical training assistant, in: 2017 International Conference on Information and Communication Technologies (ICICT), IEEE, 2017,pp. 191–196.

[14] T. Amiel, T. C. Reeves, Design-based research and educational technology: Rethinking technology and the research agenda, Journal of educational technology & society 11 (2008) 29–40.

[15] M. De Villiers, P. Harpur, Design-based research the educational technology variant of design research: illustrated by the design of an m-learning environment, in: proceedings of the South African institute for computer scientists and information technologists conference, 2013, pp. 252–261.

[16] G. D. Chittleborough, D. F. Treagust, Why models are advantageous to learning science, Educación química 20 (2009) 12–17.