Immersive Multimodal Environments for Psychomotor Skills Training

Khaleel Asyraaf Mat Sanusi[1][0000−0001−6766−4416] and Roland Klemke[1,2][0000−0002−9268−3229]

[1] Cologne Game Lab, TH Köln, Cologne, Germany
[2] Open University of the Netherlands, Heerlen, The Netherlands

Abstract. Modern immersive multimodal technologies enable the learners to completely get immersed in various learning situations in a way that feels like experiencing an authentic learning environment. These environments also allow the collection of multimodal data, which can be used with artificial intelligence to further improve the immersion and learning outcomes. The use of artificial intelligence has been widely explored for the interpretation of multimodal data collected from multiple sensors, thus giving insights to support learners' performance by providing personalised feedback. In this paper, we present a conceptual approach for creating immersive learning environments, integrated with multi-sensor setup to help learners improve their psychomotor skills in a remote setting.

Keywords: immersive learning technologies · multimodal learning · sensor devices · artificial intelligence · psychomotor training

1 Introduction

The global pandemic event of Covid-19 has shown that many learning and teaching activities can be performed in a remote manner, 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 automates them [6]. The learning model for training psychomotor training needs to be well structured, consisting of well-defined instructions and feedback that can be provided during the training. This ensures the development of such skills in a correct manner while ensuring consistent progress during the iterative practice sessions. As such, constant feedback from the teacher is essential for the learner to avoid developing improper techniques during training and ensuring the desired goal can be achieved in a shorter duration. However, a remote setting for psychomotor training makes the learning process tedious, ineffective, and inefficient, obstructing the beginner's progress. Therefore, the remote setting for psychomotor skill training affects both the learners and the teachers alike.

Nonetheless, educational researchers in the field of technology enhanced learning and artificial intelligence (AI) are continuously working to enhance psychomotor training by embedding multiple sensor technologies and machine learning models to facilitate multimodal learning experiences [2, 5, 14]. Also, technologies such as virtual reality (VR) and augmented reality (AR) are increasingly used to provide learners a more immersive feeling in an authentic setting [3, 10, 13]. The use of these combined technologies introduces new technological affordances that can be leveraged in the field of psychomotor learning, especially in a remote manner. However, experts or teachers may still be skeptical of these technologies, especially AI, since they are relatively new and not widely adopted.

Therefore, in this paper, we present a conceptual approach for creating immersive and interactive multimodal training environments using sensor and AI technologies to deliver instructions and feedback to learners in an authentic setting. The development of this prototype will be an early and significant step towards combining such technologies in an authentic setting for collecting multimodal data and giving immediate feedback for learners to improve their psychomotor skills.

2 Related Work

2.1 Sensors in Psychomotor Learning

Sensors are emerging technologies that are increasingly becoming more portable and smarter, enabling efficient methods for the acquisition of performance data, allowing effective monitoring and intervention. Sensors have been explored in the learning domain for their affordances to provide support to the learner. Schneider et al. [16] analysed 82 prototypes found in literature studies based on Bloom's taxonomy of learning domains. They explored how the prototypes can be used to provide formative assessment, especially as a feedback tool. The analysis revealed that sensor-based learning applications can assist learning scenarios in several areas, including the psychomotor domain. This research recommends researchers and educators to consider sensor-based platforms as dependable learning tools to reduce the workload of human teachers or mentors and, therefore, contribute to the solution of many current educational challenges.

In the psychomotor domain, motion sensors such as accelerometers and gyroscopes are predominantly used for the acquisition of motion data to recognize human activities. Pereira dos Santos et al. [15] emphasize the importance of combining motion sensors to obtain higher accuracy in detecting not only simple but complex activities as well. Combining these sensors enables the collection of multimodal data and, therefore, a more accurate representation of the learning process [2]. The collection of multimodal data can be done by using various types of sensors such as wearable sensors, depth cameras, Internet of Things devices, computer logs, etc. Moreover, these sensors are typically synchronized with human activities because one can analyse historical evidence of learning activities and the description of the learning process [11]. Limbu et al. [12] designed a system combining Pen sensors in Microsoft Surface and EMG sensors in Myo armband for learners to practice their calligraphy skills deliberately. It allows the calligraphy teacher to create an expert model, in which later the learners can use to practice while receiving guidance and feedback. Similarly, Schneider et al. [17] also designed a system to support the development of complex non-verbal communication skills for public speaking. The system uses the Microsoft Kinect v2 depth camera sensor that is placed in front of the learner to track the skeletal joints of the learner's body, along with a microphone, while presenting and provide real-time feedback based on common public speaking mistakes such as facial expressions, body posture, voice volume, gestures, and pauses.

To better understand learners' psychomotor performance, AI, more precisely machine learning approaches, has been explored. Educational researchers are increasingly using machine learning models to classify the activities based on the collected multimodal data. For instance, Di Mitri et al. [5] investigated to what extent multimodal data and Neural Networks can be used for learning Cardiopulmonary Resuscitation skills by utilising a multi-sensor system consisting of a Kinect for tracking body position and a Myo armband for collecting electromyogram information. Another example, Mat Sanusi [14] applied the same framework as the previous author but in a table tennis domain consisting of more dynamic movements and complex techniques. The author experimented with smartphone sensors (accelerometer and gyroscope) for motion data and a Kinect to detect forehand strokes during training. Both study results show a high classification rate of the activities when combining the sensors, emphasizing the importance of the multimodal approach in classifying complex psychomotor activities.

2.2 Immersive Learning Technologies

Immersive learning technologies, such as Augmented reality (AR) and Virtual Reality (VR), are poised to be the next educational technology medium for effective education, even in remote scenarios. Immersive learning technologies place individuals into an interactive training or learning environment, either virtually or physically, to replicate/enrich the authentic learning context of a specific skill. This provides a ubiquitous opportunity for experiential learning at any time and place. These immersive technologies can further accelerate learning experiences by providing just-in-time instructions and feedback. For instance, Chan et al. [3] developed a dance training system based on VR integrated with motion capture technology. The system projects a virtual teacher on the wall screen and demonstrates specific movements, allowing the learners to imitate. When mistakes are detected by the motion capture system, the virtual teacher provides visual immediate feedback by suggesting improvements on particular limbs at particular moves. In another example, Kyan et al. [10] implemented a VR ballet dance training system that tracks the learner's skeletal joints using the Kinect depth camera sensor. This system applies a similar concept to the "magic mirror" approach in which the virtual character projected on the wall screen moves accordingly to the learners, enabling them to reflect on their performance in real-time.

However, the use of visual feedback is commonly insufficient to not only the dance use case but in most psychomotor activities due to their complexity.

Learning psychomotor skills typically involves more complicated techniques, and learners may not be able to focus on the virtual teacher all the time. Thus, an ideal solution would be the combination of multiple modalities (e.g., visual, audio, haptic), as advocated by the multimodal approach within the immersive technologies. This can be seen in the work of Jadhav et al. [9] in which they designed a wearable soft robotic glove that provides haptic feedback in virtual environments. In their work, learning piano was used as an application case. As a result, the haptic feedback compliments the visual feedback provided from the virtual reality headset and audio feedback generated by the headphone, emphasizing the need for the multimodal approach in psychomotor learning.

While multimodal technologies provide affordances for immersion, designing a learning experience that is immersive mostly depends on the design of the learning activity. For example, Herrington et al. [8] stressed that the learning environment and task create the conditions for "True" immersion. Their study found that the task itself is the key element of immersion and engagement in cognitive (synonym to psychomotor) learning, which plays an important role in motor skills development. Similarly, it can be argued that problems presented to the student for learning should be real or that simulation should have ultra-realistic physical similarity to an actual context. Therefore, it can also be argued that the instructions and feedback provided by the learning environment should be realistic for the beginners to perform the tasks in a correct manner; for example, actual human teachers like/personalised feedback to create immersive learning experiences. Such instructions and feedback that are able to match a real human expert can be provided by using AI. In our work, we aim to create a "truly" immersive and interactive training environment with the help of immersive multimodal learning technologies to train psychomotor skills by providing human-teacher-like instructions and feedback.

3 Research Aim

Our main objective for this research is to create an immersive and interactive training environment using immersive multimodal learning technologies. This learning environment will integrate the multi-sensor setup, enabling the collection of multimodal data and, thereof, machine learning analysis to help learners improve their psychomotor skills. Therefore, the following research questions have been constructed:

RQ1 **How can we create an immersive and information rich (remote/self-learning) training environment for psychomotor skills that delivers effective instructions and meaningful feedback to the learner?**

RQ2 **Can the new immersive training environment improve existing learning methods in remote or self-directed education for psychomotor skills?**

RQ3 **How can human teachers and experts contribute to and benefit from the training environment?**

4 Research Methodologies

While immersive technologies are becoming more prominent and progressively used in the context of psychomotor training, exploiting them to create a virtual training environment has no assurance in supporting learners effectively. The intended outcome of the project is a prototypical solution for psychomotor learning that is immersive, along with a systematic process of designing that solution. Therefore, it is important for this research to follow a methodological approach for designing, developing, and testing such a prototype. Hence, we conduct our research in the Design-based Research (DBR) approach, which is a common iterative methodological approach for prototypical solutions [1] (see Fig. 1).

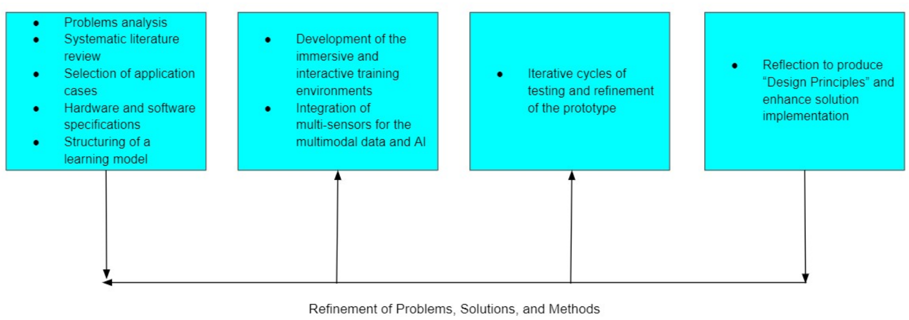

Fig. 1. The DBR approach in the context of this paper [1].

The DBR methodological approach is divided into four phases. For the first phase, it is ideal to start by analysing the problems in the domain of immersive multimodal environments that need addressing. A systematic literature review will then be followed to identify and synthesize related work publications. The next iteration will be the selection of application cases, in which each case includes several psychomotor tasks. Further, we specify the hardware and software components that are appropriate for the selected application cases. In the last iteration of this phase, a learning model, which comprises instructions and feedback, should be constructed.

The next phase should be the implementation of the training environment. A structured learning model from the previous phase should be included to create immersion for learners. Next, to further improve the immersion and learning outcomes, the training environment will be integrated with multi-sensor setup for the collection of multimodal data and AI for the machine learning analysis to provide feedback to learners based on their performance.

Subsequently, in the next phase, a user test involving the teachers/experts should be conducted to reveal essential aspects of how the prototype can be improved. Additionally, questionnaires and surveys are helpful to provide a general idea of how users perceive the interaction between the prototype. The refinement of the prototype should then be followed to ensure that the prototype is ready to be tested with the learners.

Lastly, 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 the changes have been made. Through this documentation, it can be helpful for other researchers and designers who are interested in those findings, can examine them in relation to their context and needs.

5 Conclusion

In this paper, we presented a conceptual approach for creating immersive and interactive multimodal training environments using sensors and AI to deliver meaningful instructions and effective feedback to learners in an authentic setting. The development of this prototype will be an essential step towards combining such technologies in an authentic setting for collecting multimodal data and providing immediate feedback for learners to improve their psychomotor skills remotely, thereby further enhancing the immersive learning experiences.

References

1. Anderson, T., Shattuck, J. (2012). Design-based research: A decade of progress in education research?. Educational researcher, 41(1), 16-25. https://doi.org/https://doi.org/10.3102/0013189X11428813
2. Blikstein, P., Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. Journal of Learning Analytics, 3(2), 220-238. https://doi.org/https://doi.org/10.18608/jla.2016.32.11
3. Chan, J. C., Leung, H., Tang, J. K., Komura, T. (2010). A virtual reality dance training system using motion capture technology. IEEE transactions on learning technologies, 4(2), 187-195. https://doi.org/10.1109/TLT.2010.27
4. Cratty, B. J. and Noble, . Clyde Everett (2016, April 1). Psychomotor learning. Encyclopedia Britannica. https://www.britannica.com/science/psychomotor-learning. Last accessed 10 March 2021
5. Di Mitri, D., Schneider, J., Trebing, K., Sopka, S., Specht, M., Drachsler, H. (2020, July). Real-Time Multimodal Feedback with the CPR Tutor. In International Conference on Artificial Intelligence in Education (pp. 141-152). Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-030-52237-$7_1$2
6. Feder, K. P., Majnemer, A. (2007). Handwriting development, competency, and intervention. Developmental Medicine Child Neurology, 49(4), 312-317. https://doi.org/https://doi.org/10.1111/j.1469-8749.2007.00312.x

7. Hattie, J., Timperley, H. (2007). The power of feedback. Review of educational research, 77(1), 81-112. https://doi.org/https://doi.org/10.3102/003465430298487
8. Herrington, J., Reeves, T. C., Oliver, R. (2007). Immersive learning technologies: Realism and online authentic learning. Journal of Computing in Higher Education, 19(1), 80-99.
9. Jadhav, S., Kannanda, V., Kang, B., Tolley, M. T., Schulze, J. P. (2017). Soft robotic glove for kinesthetic haptic feedback in virtual reality environments. Electronic Imaging, 2017(3), 19-24. https://doi.org/https://doi.org/10.2352/ISSN.2470-1173.2017.3.ERVR-102
10. Kyan, M., Sun, G., Li, H., Zhong, L., Muneesawang, P., Dong, N., Elder, B., Guan, L. (2015). An approach to ballet dance training through ms kinect and visualization in a cave virtual reality environment. ACM Transactions on Intelligent Systems and Technology (TIST), 6(2), 1-37. https://doi.org/https://doi.org/10.1145/2735951
11. Lahat, D., Adali, T., Jutten, C. (2015). Multimodal data fusion: an overview of methods, challenges, and prospects. Proceedings of the IEEE, 103(9), 1449-1477. https://doi.org/10.1109/JPROC.2015.2460697
12. Limbu, B. H., Jarodzka, H., Klemke, R., Specht, M. (2019). Can You Ink While You Blink? Assessing Mental Effort in a Sensor-Based Calligraphy Trainer. Sensors, 19(14), 3244. https://doi.org/https://doi.org/10.3390/s19143244
13. Limbu, B., Vovk, A., Jarodzka, H., Klemke, R., Wild, F., Specht, M. (2019, September). WEKIT. one: a sensor-based augmented reality system for experience capture and re-enactment. In European Conference on Technology Enhanced Learning (pp. 158-171). Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-030-29736-$7_1$2
14. Mat Sanusi, K. A. (2020). Stroke in Pong like Ma Long: Forehand Strokes Detection for Automatic Feedback Generation in Table Tennis Performance based on Multimodal Data and Neural Networks. Master Thesis. University of Saarland, Germany.
15. Dias Pereira dos Santos, A., Yacef, K., Martinez-Maldonado, R. (2017, July). Let's dance: how to build a user model for dance students using wearable technology. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization (pp. 183-191). https://doi.org/https://doi.org/10.1145/3079628.3079673
16. Schneider, J., Börner, D., Van Rosmalen, P. and Specht, M., (2015). Augmenting the senses: a review on sensor-based learning support. Sensors, 15(2), pp.4097-4133. https://doi.org/https://doi.org/10.3390/s150204097
17. Schneider, J., Börner, D., Van Rosmalen, P., Specht, M. (2016). Can you help me with my pitch? Studying a tool for real-time automated feedback. IEEE Transactions on Learning Technologies, 9(4), 318-327. https://doi.org/10.1109/TLT.2016.2627043