Evaluating an immersive learning toolkit for training psychomotor skills in the fields of human-robot interaction and dance.

Khaleel Asyraaf Mat Sanusi[1,2], Daniel Majonica[1,2], Patrick Handwerk[1], Suhana Biswas[3] and Roland Klemke[1,2]

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

Abstract

Developing immersive learning environments for psychomotor skills training typically demands significant time and effort. Immersive training toolkits enable the development of virtual learning environments, incorporating interactive game elements and learning materials to enrich the learning process. Such toolkits have found extensive use in classroom settings; however, their application in the psychomotor domain remains largely unexplored. This paper conducts an evaluation of IMPECT, a training toolkit designed for teaching psychomotor skills within immersive learning environments, focusing on two distinct scenarios: human-robot interaction and dancing. The study gathers survey data to assess the system's usability and incorporates participant suggestions, which are subsequently analyzed and discussed. The initial study results demonstrate the training toolkit's potential applicability across diverse psychomotor domains.

Keywords

immersive learning environments, toolkit, human-robot interaction, dance, system usability

1. Introduction

Achieving proficiency in psychomotor skills necessitates deliberate practice and technique, typically taught within a physical context. These skills demand physical execution, often through repetitive practice, to the point where the learner's muscle memory becomes trained, automating the muscle movements and enabling smoother, more natural execution [1]. That being said, the presence of teachers is essential for instructions and feedback provision. Such learning components ensure that movement errors can be promptly identified and corrected, leading to improved performances [2].

Immersive learning systems are educational tools that are typically supported by immersive technologies such as virtual reality (VR), augmented reality (AR), and mixed reality (MR),

MILeS 23: The Third International Workshop on Multimodal Immersive Learning Systems, September 4, 2023, Aveiro, Portugal
ks@colognegamelab.de (K. A. M. Sanusi); dm@colognegamelab.de (D. Majonica); roland.klemke@ou.nl (R. Klemke)
0000-0001-6766-4416 (K. A. M. Sanusi); 0000-0003-4792-0472 (D. Majonica); 0000-0002-9268-3229 (R. Klemke)

allowing the creation of immersive learning environments (ILEs) to provide realistic teaching and learning experiences in a virtual world [3, 4]. In addition, sensor technologies are used for the collection of multimodal data to track learners' behavior and performance, further aiming to improve the learning outcome [5]. By exploiting the multimodal data, instruction and feedback components can be provided in a multimodal manner (e.g., visual, auditory, haptic), conveying more and richer information to the learner [6].

An immersive learning toolkit refers to a set of resources, software, and hardware designed to facilitate the creation, implementation, and management of learning sessions or educational materials in virtual worlds. It includes various resources like learning materials, exercises, assessments, and game-world interactive elements, further enhancing the learning experience [7]. In recent years, immersive learning toolkits have been vastly developed and widely used for classroom activities [8]. However, such toolkits are still yet to be explored in the psychomotor domain for skills training.

This paper presents an evaluation of the Immersive Multimodal Psychomotor Environments for Competence Training (IMPECT) toolkit in two specific application areas within the psychomotor domain: human-robot interaction (HRI) and dance. The assessment encompasses the current state of prototypical development, usability scores, and recommendations for enhancements drawn from the survey data.

This article is structured as follows. Section 2 describes the proposed toolkit. Section 3 explains the methodology and shows the results of the collected data. We summarize and discuss the findings of the study in Sections 4 and 5, respectively. Lastly, limitations and future work are described in Section 6.

2. Immersive Multimodal Psychomotor Environments for Competence Training (IMPECT)

IMPECT represents a versatile training toolkit that facilitates the development of ILEs, leveraging sensors and immersive technologies to cater to various psychomotor skills training needs. Acknowledging that different domains might call for specific technologies and learning materials, this toolkit empowers developers, researchers, and educators to create customized ILEs tailored to their psychomotor skills training objectives.

2.1. IMPECT Clients

In the context of this paper, two client applications were developed using Unity, exploring the domains of HRI and Dance. Both are designed for visualizing and providing feedback from the IMPECT-server (see Section 2.2). In order to do so, an established connection is made between both clients and the server to ensure that feedback can be triggered by human experts. Since the development of skills varies between the two cases, different technologies are utilized. The following subsections introduce the two clients, namely IMPECT-HRI and IMPECT-Dance.

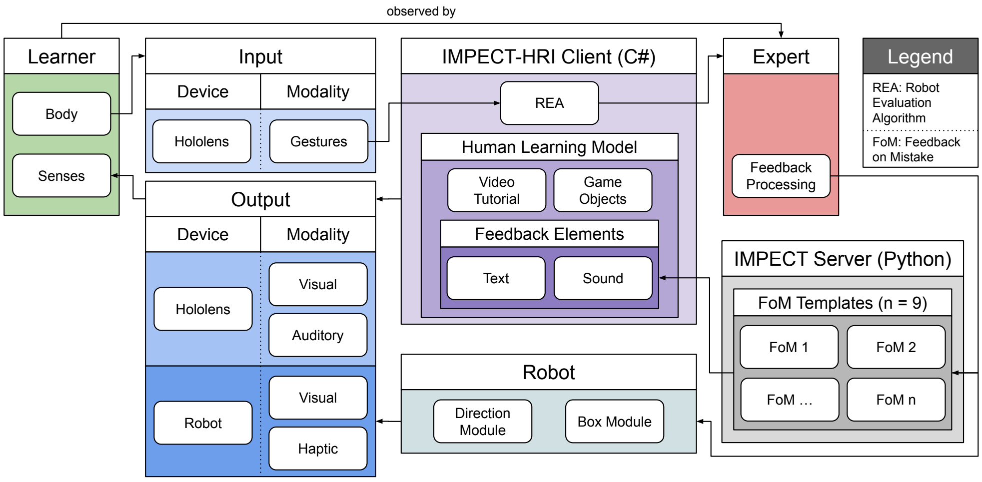


Figure 1: IMPECT-HRI architecture.

2.1.1. IMPECT-HRI

The IMPECT-HRI tool is an AR application created using IMPECT. As sensors, it uses the gesture detection of the Hololens. Inside the tool, the user can autonomously interact with the UI components that are presented in the AR environment. Whenever an interaction through the system happens, it will get logged and analyzed. The Robot Evaluation Algorithm (REA) on the AR hardware takes the current state of the environment and the interaction with the system into account and calculates a decision. This decision runs through the expert who also observes the learner doing the task. If the learner did a mistake or something unexpected happens, the expert can react and send feedback to the IMPECT server. On the server, multiple pre-determined "Feedback on the Mistake" (FoM) can be selected. The calculated decision by the REA gets transported to the physical robot. This physical robot takes in the decision and visually displays its decision to the learner. The feedback on the other hand gets sent through the AR environment and gets displayed in the form of visual and auditory feedback to the learner. This feedback is displayed either through text, sound, or both. This whole process can be seen summarized in Figure 1.

2.1.2. IMPECT-Dance

The IMPECT-Dance tool is a desktop-based application that utilizes the Kinect's depth camera sensor, focusing on basic dance movements. Additional feedback and UI components were implemented in the tool, extending the previous version that concentrate on basic exercise routines [9]. For visualizing the learning environment with its feedback components, an external screen is used.

Figure 2 shows the system architecture of the IMPECT-Dance tool. The visualization of the learners is captured by the Kinect camera, mirroring them and allowing them to see themselves when performing the dance movements. The human learning elements comprise instructional

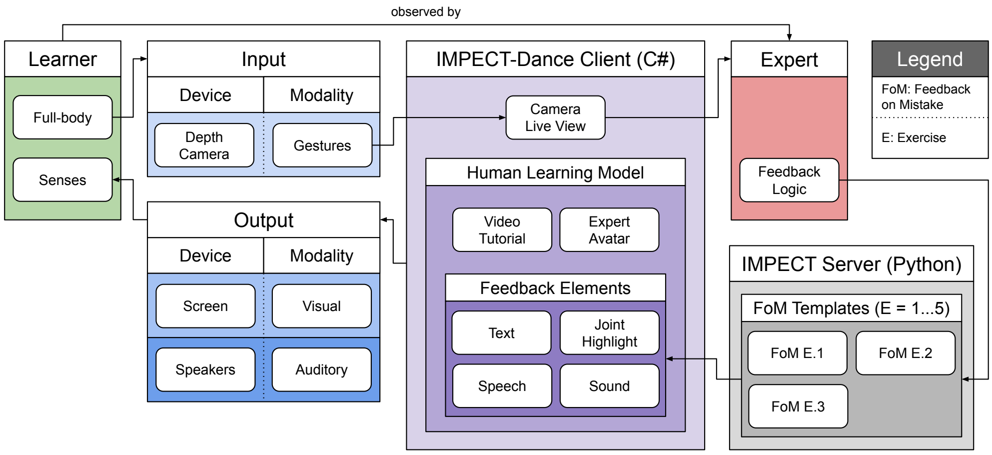


Figure 2: IMPECT-Dance architecture.

objects and feedback objects. The instruction objects consist of a Video Tutorial of the expert performing a full-dance routine which is segregated into five different exercises. The presence of the Expert Avatar aims to assist the learners by performing the animation of the five exercises procedurally. The animations were pre-recorded by the expert in the video with the use of a motion capture system for recording the dance movements.

The **feedback elements** consist of both visual (text and highlighted-joint) and auditory (sound and speech) modalities. For visual feedback, the specific joint/s of the avatar will be highlighted in red/orange color, which corresponds to the mistake. Additionally, text feedback is given in a way to guide the learner in correcting the mistake. For auditory feedback, a 90 BPM metronome sound is played throughout the session to assist the learner in following the beat. Furthermore, speech feedback is played when a mistake is detected, corresponding to the same message as the text feedback. In the context of our paper, the expert drives the simulation by observing the movements of the learner and selecting the corresponding FoM template when a mistake is detected.

2.2. IMPECT Server

The server was implemented using Python in order to provide full control for the expert/teacher in navigating the learning session and sending FoM templates according to the mistakes that have been detected. The user interface (UI) is designed to be straightforward for the teacher, consisting of tabs and buttons that would steer the sub-sessions and trigger feedback, respectively. Each application comprises instructions and FoM templates that can be selected and sent to the respective client. For each FoM template, several feedback types (UI or game-world objects) and modalities (visual and/or auditory) are clustered, allowing the teacher to prompt the learner by choosing the specific FoM that corresponds to the detected mistake.

3. Methods

3.1. Research Design

Two distinct events were utilized to evaluate both IMPECT tools, with a total of 29 participants attending the two sessions. To minimize any potential disruptions and uphold the element of surprise for each study, the setup for both experiments took place at separate locations. Following a brief introduction, the participants were asked to read the information letter and provide their consent by signing the form.

3.1.1. HRI

Before interacting with the system, the learner first had to be introduced to the game which is being played cooperatively with the robot. It involves a physical robot, an AR environment, and custom playing cards. Instructions are provided via video and feedback is provided in the form of text and sound. The game rules are as follows:

- Only go through rooms that are connected by a door.
- In order to move to a different room; first, click the virtual card, then, turn the physical card
- In robot rooms; first, click the virtual card, second, wait for the robot's decision, third, make your final decision, then, turn the physical card
- Find the keycard inside a hidden box (hidden object card) to open the last door to the engine room.
- A maximum of 15 turns to traverse the maze.

The protocol for communicating the game rules to participants within the AR ILE employs instructional videos comprising both visual and auditory components. This video is being played at the beginning of the research study. The game progression can be summarized as follows: 1) The player first decides which room to enter on the virtual grid. This choice is made within the AR environment by selecting a room. 2) Some rooms are labeled as "robot rooms." Within these rooms, the robot's decision-making algorithm (REA) makes decisions based on the ongoing game situation. Once the REA decides, the player must reconfirm their choice of room within the virtual grid. 3) Once the virtual room choice is confirmed, the player is required to flip the corresponding physical card. An essential aspect of completing the game involves locating a hidden box, symbolized by a concealed card somewhere within the maze. Successfully finding and using this hidden card is necessary to conclude the game.

The REA uses a path-finding algorithm to navigate the maze. However, there are modifications to the algorithm for different scenarios, such as when the learner has turned left, REA always agrees with the learner's choice. On the last decision, REA deliberately disagrees with the learner's choice to observe their reactions to different outcomes.

3.1.2. Dance

The full-dance routine was segregated into five different exercises (Exercises 1 to 5), which the learner needs to complete in a procedural manner. The performance is observed by a sports-

Prototype	SUS
IMPECT-HRI	55.94
IMPECT-Dance	74.88
Total SUS	69.66

Table 1
SUS score

scientific expert who critically rated the users' motion executions and provided immediate relevant feedback (either praising or corrective) by selecting the FoM template from the IMPECT server accordingly to the mistake that was detected. We listed three common beginner mistakes for each exercise (15 in total) in the server, classifying them as mistake IDs. Each FoM template corresponds to the designated mistake ID, consisting of several feedback types and modalities. For instance, "FoM 1.2" happens when the learner's elbows are too low, therefore, elbows need to be highlighted (visual), text (visual), and speech (auditory) feedback are given to the learner.

The study comprises two sessions: learning and challenge. The learning session allows the learner to perform the movements procedurally and receives feedback when mistakes are detected. The challenge session, however, requires the learner to perform the full-dance routine without receiving any feedback. In total, 21 people participated in the dance study.

3.2. Instruments

To assess the usability of both prototypes, the System Usability Scale (SUS) was employed. This 5-point Likert scale offers a comprehensive evaluation of subjective user assessments, encompassing effectiveness, efficiency, and satisfaction aspects of the systems [10]. Moreover, the participants were presented with five open-ended questions, inviting their valuable input and suggestions for future improvements to be incorporated in subsequent iterations of the prototypes.

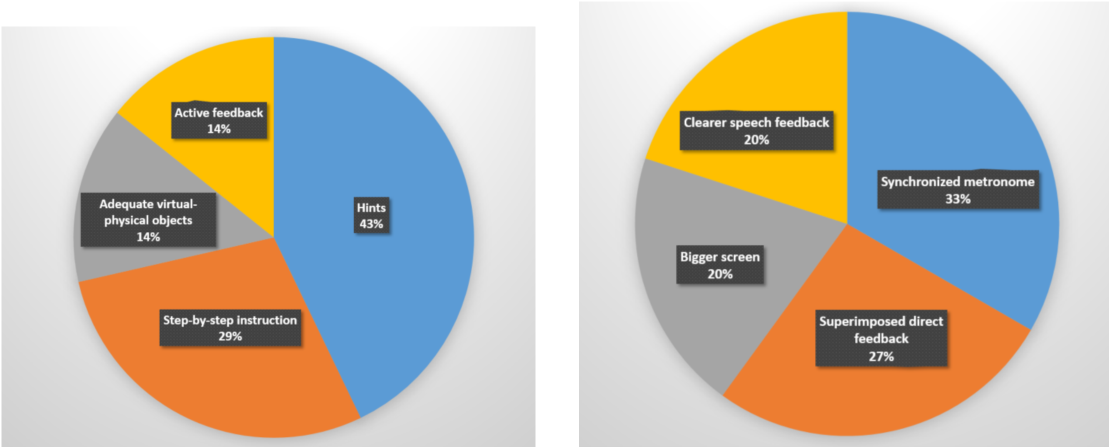
4. Results

System's Usability Scale

In Table 1, we observed that the SUS scores for IMPECT-HRI and IMPECT-Dance are 55.95 and 74.88, respectively. Overall, the SUS score for both prototypes has a mean score of **69.66**, which is slightly above the average score (68).

Suggestions and Improvements

Figure 3 shows the three most suggested components by the participants. In the case of HRI (Figure 3a), "Hints" (43%) is considered to be the most crucial component that needs to be implemented to the IMPECT-HRI tool, followed by "Step-by-step instruction" (23%), "Active feedback" (14%), and "Adequate virtual-physical objects" (14%). For the dance case (Figure 3b), "Synchronized metronome" (33%) was the most suggested component for the IMPECT-Dance



(a) IMPECT-HRI (b) IMPECT-Dance

Figure 3: Suggestions for both IMPECT tools.

tool, followed by "Superimposed direct feedback" (27%), "Clearer speech feedback" (20%), and "Bigger screen" (20%).

5. Discussion and Conclusion

Based on the collected data, we found that both prototypes received a relatively positive usability rating, with a SUS score of 69.66. Although this score is slightly higher than the threshold, it indicates that participants generally perceived the prototypes as reasonably usable and user-friendly. Notably, in the case of the dance prototype, it scored higher on the SUS scale, which could be attributed to the simplicity of tasks involved. The dance movements integrated into the system were repetitive and easier to perform compared to the HRI case, which required participants to engage in more complex tasks. The familiarity factor also played a role. Learning dance skills is more common and accessible compared to the relatively novel concept of interacting or collaborating with robots. Additionally, as the prototypes involved mixed-reality glasses and robot interaction, which are not widely accessible in society, many participants might not have been familiar with the gestures needed to navigate the learning environment, contributing to the perceived complexity of handling the prototype.

For IMPECT-HRI, the survey highlighted the need for a hints system as the prototype's instructions before the card game were unclear, causing delays in the current session. Furthermore, participants suggested video instructions delivered in a step-by-step manner. Due to the passiveness and low level of feedback immediacy, the participants proposed the feedback to be more active and immediate. Interestingly, some participants pointed out a weak UI/UX connection between the virtual and physical elements, urging improvements for better integration and meaningful learning outcomes.

Regarding IMPECT-Dance, the most crucial suggestion was to address the synchronization issue between the metronome and the dancing avatar, which became increasingly unaligned during longer sessions. Participants also found the coexistence of both the live view and the expert avatar confusing, leading to the recommendation of superimposed feedback on the live avatar, based on the participant's movements. The audio feedback was perceived to be slightly inaudible, prompting the implementation of speech feedback to make it more easily audible to participants. Lastly, participants expressed the need for a larger, higher-quality screen to enhance the visualization of the environment during training.

This paper presents a comprehensive evaluation of the IMPECT training toolkit, focusing on two application areas within the psychomotor domain: human-robot interaction (HRI) and dance. The assessment encompasses three main aspects: the current state of prototypical development, the usability scores acquired, and the recommendations for enhancements based on the survey data. The questionnaire results served as valuable feedback to the researchers and developers, guiding potential user-specific adaptations and future improvements.

In summary, this paper offers both theoretical and practical implementations. From a theoretical standpoint, it introduces the concept of FoM templates, comprising diverse feedback types that can be utilized for skills training and expanded with additional components tailored to specific use cases. On the practical side, the toolkit's application extends to various psychomotor domains, with the possibility of further exploration through the addition of more use cases to enhance its scalability and flexibility.

6. Limitations and Future Work

During the study, several shortcomings came to light. Firstly, we faced limitations in experimenting with all workshop participants due to the time constraints of the workshop itself. Furthermore, some participants encountered difficulties with the visualization of the environment because of inadequate brightness from the projector. The technology utilized also had its limitations, affecting the accessibility of the environment. Additionally, a participant with poor eyesight experienced challenges in interacting with the environment effectively.

In future endeavors, there is room for enhancing both prototypes by incorporating a wider range of feedback elements, including various types and modalities, to foster more effective and meaningful learning outcomes. Consideration can also be given to integrating artificial intelligence technology, enabling performance analysis and feedback provision in an automated manner. Furthermore, it is worth exploring additional psychomotor use cases alongside their respective technologies and FoM templates. This approach aligns with the overarching concept of the IMPECT training toolkit, designed to cater to diverse psychomotor domains and offer a comprehensive solution for training and competence development.

7. Acknowledgments

This work was funded by the German Federal Ministry of Education and Research (BMBF) within the project "MILKI-PSY"[1] under the project ID: 16DHB4015.

[1] https://milki-psy.de/

References

[1] R. W. Barnes, Surgical handicraft: Teaching and learning surgical skills, The American Journal of Surgery 153 (1987) 422–427. doi:10.1016/0002-9610(87)90783-5.

[2] F. W. Otte, K. Davids, S.-K. Millar, S. Klatt, When and how to provide feedback and instructions to athletes?—how sport psychology and pedagogy insights can improve coaching interventions to enhance self-regulation in training, Frontiers in Psychology 11 (2020) 1444.

[3] K. A. M. Sanusi, D. Iren, R. Klemke, Experts' evaluation of a proposed taxonomy for immersive learning systems, in: International Conference on Games and Learning Alliance, Springer, 2022, pp. 247–257.

[4] M. Geisen, S. Klatt, Real-time feedback using extended reality: A current overview and further integration into sports, International Journal of Sports Science & Coaching 17 (2022) 1178–1194.

[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, volume 12163 of *Lecture Notes in Computer Science*, Springer International Publishing, Cham, 2020, pp. 141–152. doi:10.1007/978-3-030-52237-7{\textunderscore}12.

[6] D. Di Mitri, J. Schneider, M. Specht, H. Drachsler, From signals to knowledge: A conceptual model for multimodal learning analytics, Journal of Computer Assisted Learning 34 (2018) 338–349.

[7] H.-K. Jee, S. Lim, J. Youn, J. Lee, An augmented reality-based authoring tool for e-learning applications, Multimedia Tools and Applications 68 (2014) 225–235.

[8] A. Dengel, M. Z. Iqbal, S. Grafe, E. Mangina, A review on augmented reality authoring toolkits for education, Frontiers in Virtual Reality 3 (2022) 798032.

[9] K. A. M. Sanusi, M. Slupczynski, M. Geisen, D. Iren, R. Klamma, S. Klatt, R. Klemke, Impect-sports: Using an immersive learning system to facilitate the psychomotor skills acquisition process., in: Seventeenth European Conference on Technology Enhanced Learning. Proceedings of the Second International Workshop on Multimodal Immersive Learning Systems., 2022.

[10] J. Brooke, et al., Sus-a quick and dirty usability scale, Usability evaluation in industry 189 (1996) 4–7.