



**HAL**  
open science

# ImCasting: Nonverbal Behaviour Reinforcement Learning of Virtual Humans through Adaptive Immersive Game

Miriam Punzi, Nicolas Ladeveze, Huyen Nguyen, Brian Ravenet

► **To cite this version:**

Miriam Punzi, Nicolas Ladeveze, Huyen Nguyen, Brian Ravenet. ImCasting: Nonverbal Behaviour Reinforcement Learning of Virtual Humans through Adaptive Immersive Game. IUI '22: 27th International Conference on Intelligent User Interfaces, Mar 2022, Helsinki, Finland. pp.62-65, 10.1145/3490100.3516475 . hal-03747416

**HAL Id: hal-03747416**

**<https://hal.science/hal-03747416v1>**

Submitted on 8 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ImCasting: Nonverbal Behaviour Reinforcement Learning of Virtual Humans through Adaptive Immersive Game

MIRIAM PUNZI, University Paris-Saclay, CNRS, LISN, France

NICOLAS LADEVEZE, University Paris-Saclay, CNRS, LISN, ASARD Team, France

HUYEN NGUYEN, University Paris-Saclay, CNRS, LISN, VENISE Team, France

BRIAN RAVENET, University Paris-Saclay, CNRS, LISN, CPU Team, France



Fig. 1. The ImCasting game virtual environment: A) The main stage of the audition room shows all of the five actors standing on the trapdoors with their masks on waiting to be judged by a player as the movie director after their performance; B) The player is able to express different emotions (e.g., joy emotion in this figure) to give an example to the system at the beginning of the game; and C) The skeleton of the player's avatar reproduces this upper body motion in the virtual world.

This research work puts a focus on the user experience of an alternative method to teach nonverbal behaviour to Embodied Conversational Agents in immersive environments. We overcome the limitations of the existing approaches by proposing an adaptive Virtual Reality game, called ImCasting, in which the player takes an active role in improving the learning models of the agents. Specifically, we based our approach on the Human-in-the-loop framework with human preferences to teach the nonverbal behaviour to the agents through the system. We introduce game mechanisms built around all the tasks of this Machine Learning framework, designing how a human should interact within this framework in real-time. The study explores how a game interaction in an immersive environment can improve the user experience in performing this interactive task, sharing the same space with the learning agents. In particular, we focus on the involvement of the players as well as the usability of the system. We conducted a preliminary evaluation to compare the design of our system with a baseline system which does not use any game mechanisms in teaching the nonverbal behaviour to virtual agents. Results suggest that our design concept and the game story are more engaging, increasing the satisfaction usability factor perceived by the participants.

## ACM Reference Format:

Miriam Punzi, Nicolas Ladeveze, Huyen Nguyen, and Brian Ravenet. 2022. ImCasting: Nonverbal Behaviour Reinforcement Learning of Virtual Humans through Adaptive Immersive Game. 1, 1 (March 2022), 5 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

Authors' addresses: Miriam Punzi, University Paris-Saclay, CNRS, LISN, Orsay, France, [miriam.punzi@limsi.fr](mailto:miriam.punzi@limsi.fr); Nicolas Ladeveze, University Paris-Saclay, CNRS, LISN, ASARD Team, Orsay, France, [nicolas.ladeveze@lisn.fr](mailto:nicolas.ladeveze@lisn.fr); Huyen Nguyen, University Paris-Saclay, CNRS, LISN, VENISE Team, Orsay, France, [huyen.nguyen@universite-paris-saclay.fr](mailto:huyen.nguyen@universite-paris-saclay.fr); Brian Ravenet, University Paris-Saclay, CNRS, LISN, CPU Team, Orsay, France, [brian.ravenet@limsi.fr](mailto:brian.ravenet@limsi.fr).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2022 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM

1

## 1 INTRODUCTION

Embodied Conversational Agents (ECAs) are interactive computer systems that replicate the same communicative modalities as humans, verbal and nonverbal, through the use of an anthropomorphic character [9]. These interfaces are becoming more popular and effective in various applications, from customer service [15] to personalized health coaching [18, 22] and pedagogical tutoring [20]. In order to achieve human-like behaviors, agents need to have an interactive loop allowing them to receive as inputs behaviours and feedback, like a reward mechanism, from the users and to produce variations in their own behaviours according to this input. This is actually close to how Interactive Reinforcement Learning works [11] and similar to how humans learn to behave as well [5]. In the Interactive Reinforcement Learning approach, the feedback is not directly provided by the environment (or by the demonstration), but by an external trainer [16]. This approach is used for trying to solve a variety of problems where the traditional reward for the reinforcement learning algorithm is hard to define [6].

However, such an approach requires users to remain engaged in the task. An immersive environment might alleviate this constraint [13]. In [10], the authors reported that Taichi learners in an immersive environment had a better learning experience than learners in non-immersive setups. Since the task of teaching motions is complex and abstract, an immersive system provides the trainers a means of visualisation and improves the trainer experience [12]. Moreover, games and game mechanics (through the gamification process) have been used in interactive systems to create more engaging experiences as well, from promoting healthy behaviours [21] to support learning [19]. These mechanisms can alleviate the difficulty of a task by providing interactions in line with the user’s preferences (e.g., exploring, competing, etc.) [17] and have been used to design more engaging interactive learning interactions [14].

Our objective is therefore to build an ECA framework tailored for the interactive learning of its behaviours, in virtual reality (VR), thanks to feedback given explicitly by the user. Our long term objective is to build an understanding of how games and interactive learning could work together in creating adaptive and evolving interactions.

## 2 IMCASTING: THE IMMERSIVE CASTING GAME

We propose a VR adaptive game called *ImCasting*, where the user can teach nonverbal behaviour to ECAs in the context of a movie casting. The player takes the role of a movie director, who is looking for an actor to play the role of a masked character in a silent movie. The director has to hire the actor who better performs an emotion with the upper part of the body and get rid of the bad ones who try to invade their audition through different rounds of the game. Amongst the actors, some are real humans and others are virtual agents (Fig. 1.A). The characters that represent human actors play a prerecorded performance, while the virtual ones generate a performance Proximal Policy Optimization reinforcement learning algorithm implementation from the Unity ML-Agent’s plugin, in real-time.



Fig. 2. The interactive smart table.

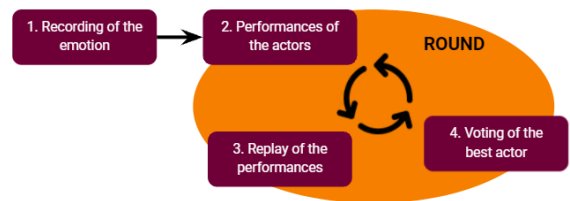


Fig. 3. The sequence of events in the loop of the game.

There are many different constraints in terms of immersive technology and the nature of the interaction between the user and virtual agents, that lead us to design considerations. For example, we limit the nonverbal behaviour training to the agents' upper-body part without considering facial expressions, and the agents have to perform the emotion of joy.

The designed game is a Seated VR [4] experience where the interaction is mainly situated in a smart table (Fig. 2). It presents a screen that shows the instructions of the story, a "YES" and a "NO" buttons used to answer the questions displayed on the screen, and some numbered buttons used to vote or to ask for a replay of a specific actor.

Fig. 3 shows the phases of the game. The agents are trained after every loop of feedback from the user as the movie director by a voting mechanism of the best performance and from the example given by the user directly.

The ideal platform would be a multi-user system in which several users interact with the agents and with each other, to enhance the user experience and to enrich the feedback. However, as the first prototype of the platform, we focus only on the single-user aspect of our system. Our main research interest is to keep the user engaged in performing the training task, investigating the appropriate user-agent interaction and the mechanism of the agents in receiving the feedback. Solving the machine learning problem of training the virtual agents is out of the scope of this work.

We developed the system using the Unity Engine and the Unity ML-Agents Toolkit [3]. Since the latter does not provide a mechanism which integrates the human feedback into the reward function, our prototype uses only the joy emotion expression given by the user as the targets to imitate for the virtual agents. We used an HTC Vive Pro headset and two controllers to record the movement of the user (Fig. 1.B). Then, the upper-body motion of the virtual agents is calculated using the Inverse Kinematic approach (IK) [1]. Each neural network of the agents learns how to move the corresponding body part by predicting the position and rotation of the targeting movement from the IK. Finally, to drive the virtual characters that represent the human actors, using Optitrack Motive, we captured several human motions of different expressions of joy which can be directly replayed in the user interface of the ImCasting game.

### 3 PRELIMINARY STUDY

We conducted a preliminary study to evaluate how our ImCasting game can help the user to be more involved in the training of the nonverbal behaviour to the virtual agents, compared to its corresponding non-gamified version. The study followed a balanced within-subject design to assess the two SYSTEM conditions. First, *ImCasting Game* condition allows the user to select the best actor from five performers, performing one by one in each round. There are two human actors who performed the motion which was prerecorded for the performers and three virtual actors whose motions were generated on the fly. Second, *No-Game* condition enables the user to compare the performance of two actors and to choose a preferred one at each round, which is inspired by the interactive reinforcement learning approach described in [11]. In order to have a learning goal, the user also records their expression of joy in this condition.

We recruited six participants (two females, four males), aged between 22 and 31 ( $m = 27$ ,  $std = 3.2$ ) for the study. After the training before each condition, the participant performed the task: first perform the motion of joy emotion, then going through rounds of selecting the best actors until they decided to stop the task. They filled out three questionnaires (System Usability Scale (SUS) [8], Godspeed Questionnaire Series (GQS) [7], and Presence [23]) and had an open-ended interview after each condition. The interview was conducted to discover opinions of the participant which are specific to each SYSTEM tested and to understand their subjective evaluation by comparing the two SYSTEMS. At the end of the experiment, the participant was also asked to rank the two conditions according to their preference. The whole experiment lasted approximately 1 hour.

From the result of the SUS questionnaire, the score of the *Game* condition is 70.42 and of the *No-Game* is 65.83. According to the interpretation of results given in [2], *Game* is considered as **Good**, while *No-Game* is considered to be

**Poor.** However, the results from the GodSpeed and Presence questionnaires show that the difference between the two conditions is minimal even if the answers are overall more positive in favor of the *Game* system (Fig. 4 left & middle).

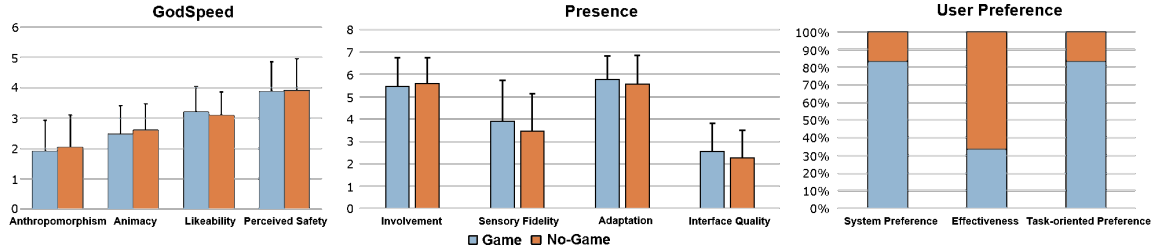


Fig. 4. The subjective rating of the participants on the two conditions (e.g., *Game* and *No-Game*) using GodSpeed Questionnaire (left) on 5-point Likert-scale, Presence (middle) on 7-point Likert-scale, and User Preference (right).

Figure 4 (right) also shows the results of the subjective preference when the participants rated the two conditions in the open-ended interview with regard to System Preference, Effectiveness, and Task-oriented Preference. Specifically, five out of six participants (83.33%) preferred the *Game* system thanks to the game adaptation idea, the decoration objects and the presence of more actors. Even if they spent more time on the task, they did not realise it because it was fun. Only one participant preferred the *No-Game* system. They thought that this system was easier to use because there were only two actors to consider, while in the *Game* condition there were too many performances to remember. This trend in preference also affected the ranking of the participants when they were asked which system they would choose to train virtual characters (e.g., Task-oriented Preference). Regarding the Effectiveness, four participants (66.67%) believed that *the No-Game* system was more effective, because it was faster and more efficient as they could focus only on one character at one time. In contrast, the other two participants (33.33%) who tested the *Game* first thought that the *Game* system was more effective because they would spend more time playing as it was more fun.

#### 4 CONCLUSION

In this work, we investigated the design of gamified interactions to support the interactive reinforcement learning paradigm in the context of training nonverbal behaviours to embodied conversational agents. Specifically, we designed a virtual reality game, ImCasting, where the player would improve the learning models of the agents through different rounds of the game. The learning objective for the agents in our study was to perform the emotion of joy with the upper body, without facial expressions. Our ultimate goal was to assess if inserting game mechanisms around the framework could elicit a more engaging interaction between a human and the framework. Such a framework could help to tailor complex adaptive interactions leveraging the social dimensions of interpersonal exchanges, leading to the development of more engaging experiences in a variety of contexts such as learning or personal coaching for instance.

The results of a preliminary study show that users were more involved and satisfied in the gamified scenario. However, the usability of the system was affected by the incapacity of the virtual agents in correctly performing the emotion of joy in the end. The fact that the system is not considering the feedback of the user yet is negatively affecting the user experience and the satisfaction factor. Our future work would be to improve the Machine Learning model in order to better align our design with our end goal, while investigating the potential of integrating more behaviors and additional tracking technologies such as face and fingers.

## REFERENCES

- [1] Inverse kinematics. <https://docs.unity3d.com/Manual/InverseKinematics.html>. Accessed: 2021-08-03.
- [2] Measuring and interpreting system usability scale (sus). <https://uiuxtrend.com/measuring-system-usability-scale-sus/>. Accessed: 2021-08-19.
- [3] ML-agents. <https://github.com/Unity-Technologies/ml-agents>. Accessed: 2021-08-17.
- [4] Seated vr. [https://xinreality.com/wiki/Seated\\_VR](https://xinreality.com/wiki/Seated_VR). Accessed: 2021-08-12.
- [5] Andrew J Arnold and Piotr Winkielman. The mimicry among us: intra-and inter-personal mechanisms of spontaneous mimicry. *Journal of Nonverbal Behavior*, 44(1):195–212, 2020.
- [6] Christian Arzate Cruz and Takeo Igarashi. A survey on interactive reinforcement learning: Design principles and open challenges. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, pages 1195–1209, 2020.
- [7] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009.
- [8] John Brooke. Sus: A retrospective. *Journal of Usability Studies*, 8(2):29–40, 2013.
- [9] Justine Cassell. Embodied conversational interface agents. *Communications of the ACM*, 43(4):70–78, 2000.
- [10] Xiaoming Chen, Zhibo Chen, Ye Li, Tianyu He, Junhui Hou, Sen Liu, and Ying He. Immertai: Immersive motion learning in vr environments. *Journal of Visual Communication and Image Representation*, 58:416–427, 2019.
- [11] Paul Christiano, Jan Leike, Tom B Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. *arXiv preprint arXiv:1706.03741*, 2017.
- [12] Chris Christou. Virtual reality in education. In *Affective, interactive and cognitive methods for e-learning design: creating an optimal education experience*, pages 228–243. IGI Global, 2010.
- [13] Julia Diemer, Georg W Alpers, Henrik M Peperkorn, Youssef Shiban, and Andreas Mühlberger. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in psychology*, 6:26, 2015.
- [14] Aviv Elor and Sri Kurniawan. Deep reinforcement learning in immersive virtual reality exergame for agent movement guidance. In *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–7. IEEE, 2020.
- [15] Ulrich Gnewuch, Stefan Morana, and Alexander Maedche. Towards designing cooperative and social conversational agents for customer service. In *ICIS*, 2017.
- [16] Vinicius G Goecks. *Human-in-the-Loop Methods for Data-Driven and Reinforcement Learning Systems*. PhD thesis, Texas University, 2020.
- [17] Ana Carolina Tomé Klock, Isabela Gasparini, Marcelo Soares Pimenta, and Juho Hamari. Tailored gamification: A review of literature. *International Journal of Human-Computer Studies*, 144:102495, 2020.
- [18] Lean L Kramer, Silke Ter Stal, Bob C Mulder, Emely de Vet, and Lex van Velsen. Developing embodied conversational agents for coaching people in a healthy lifestyle: Scoping review. *Journal of medical Internet research*, 22(2):e14058, 2020.
- [19] Jenni Majuri, Jonna Koivisto, and Juho Hamari. Gamification of education and learning: A review of empirical literature. In *Proceedings of the 2nd international GamiFIN conference, GamiFIN 2018*. CEUR-WS, 2018.
- [20] Ati Suci Dian Martha and Harry B Santoso. The design and impact of the pedagogical agent: A systematic literature review. *Journal of Educators Online*, 16(1):n1, 2019.
- [21] Lamyae Sardi, Ali Idri, and José Luis Fernández-Alemán. A systematic review of gamification in e-health. *Journal of biomedical informatics*, 71:31–48, 2017.
- [22] Silke ter Stal, Lean Leonie Kramer, Monique Tabak, Harm op den Akker, and Hermie Hermens. Design features of embodied conversational agents in ehealth: a literature review. *International Journal of Human-Computer Studies*, 138:102409, 2020.
- [23] Bob G Witmer and Michael J Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.