POLITECNICO DI TORINO
Repository ISTITUZIONALE

Investigating tangible user interaction in mixed-reality robotic games

Original
Investigating tangible user interaction in mixed-reality robotic games / Prattico', F.G., Baldo, P., Cannavo', A., Lamberti, F.. - STAMPA. - (2019), pp. 205-210. (9th IEEE International Conference on Consumer Electronics - Berlin Berlin, Germany September 8-11, 2019) [10.1109/ICCE-Berlin47944.2019.8966135].

Availability:
This version is available at: 11583/2748639 since: 2020-02-03T20:24:35Z

Publisher:
IEEE

Published
DOI:10.1109/ICCE-Berlin47944.2019.8966135

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

09 November 2026

Investigating Tangible User Interaction in Mixed-Reality Robotic Games

F. Gabriele Pratticò, Piero Baldo, Alberto Cannavò, and Fabrizio Lamberti
Politecnico di Torino, Dipartimento di Automatica e Informatica, Torino, Italy
filippogabriele.prattico@polito.it, piero.baldo@studenti.polito.it, alberto.cannavo@polito.it, fabrizio.lamberti@polito.it

***Abstract*—Among the emerging trends in Human-Robot Interaction, some of the most frequently used paradigms of interaction involve the use of Tangible User Interfaces. This is especially true also in the field of robotic gaming and, more specifically, in application domains in which commercial off-the-shelf robots and projected Mixed Reality (MR) technology are combined together. The popularity of such interfaces, also in other domains of Human-Machine Interaction, has led to an abundance in the number of gestures that can be used to perform tangible action using these interfaces. However, there are not sufficient pieces of evidence on how these different modalities can impact the user experience, in particular when interacting with a robot in a "phygital play" environment. By moving from this consideration, this paper reports on the efforts that are ongoing with the aim to investigate the impact of diverse gesture sets (which can be performed with the same physical prop) on the perception of interaction with the robotic system. It also presents preliminary insights obtained, which could be exploited to orient further research about the use of such interfaces for interaction in MR-based robotic gaming and related scenarios.**

***Index Terms*—Human-Robot Interaction, tangible interfaces, phygital play, game design, Mixed Reality, robotic game, user experience.**

I. Introduction and Background

In the form of commercial off-the-shelf (COTS) products, robots are getting more and more common at the consumer level. Service robots (vacuum cleaners, lawnmowers, etc.) and toy robots, in particular, are undoubtedly the most popular [1]. While technology is reaching maturity and the economic importance of the sector is growing, industry and academy are dedicating significant efforts in improving Human-Robot Interaction (HRI) [2] with the aim to raise acceptance of this technology and foster adoption at the consumer level.

One of the emerging trends regarding social robotics and HRI is their application to gaming. The dualistic physical and digital nature of robots disclosed interesting paradigms of interaction that were subject to numerous investigations. In this respect, a noteworthy HRI paradigm is reflected by the emerging category of so-called Physically Interactive Robotic Games (PIRGs) [3], which promote the role of robots as rational agents that may interact with players in a physical (and safe) way. PIRGs are a component of a more broader kind of gaming that was introduced in [4] and referred to as "phygital play": the key idea is that the digital pieces of information mediated throughout physical elements, robots included, should lead users to a more immersive and intense experience. The most straightforward and typical approach to this paradigm is based on the exploitation of Mixed Reality (MR) environments created by "augmenting" the physical world with digital contents. These environments are frequently implemented using spatial Augmented Reality (AR) [5], as proved by the amount of investigations conducted in this direction [4]–[14].

Due to the nature of the Phygital Robotic Games (PRG), many studies have considered leveraging Tangible User Interfaces (TUIs) to satisfy the requirements of such HRI paradigm [15]. A possible definition of TUI was given in [16] as utilizing physical representation and manipulation of digital data, offering interactive couplings of physical artefacts with computationally mediated digital information. Albeit TUIs, their variations and specializations (like graspable user interfaces [17]) were investigated by a plethora of studies in the last decade, their application in the field of PRGs can be summarized in three categories of investigations. The first category regards the usage of the robot itself as a TUI [18], in such a way that the robot is employed as a sort of active controller for direct manipulation of game-related data. The second category is about using TUIs that are entangled to digital elements which belong to the same space of interaction of the robot [12], [19]. Lastly the third category includes the usage of TUIs as a proxy to control the robot [20]. However, it can be debated whether this last category can be regarded as an actual implementation of a TUI, since no direct manipulation of representative digital contents is involved in the interaction.

Furthermore, plenty of studies were carried out to investigate the size and shape of TUIs [21], [22], as well as how a user can effectively manipulate them [23]. Notwithstanding, the large majority of these probes compare a single interaction modality w.r.t. traditional ones, as in [24]. A fairly significant amount of studies tackled the domain of gestures detection with various equipment, including smartphones, like in [25]. Smartphones, in particular, are often used a form of TUI, but the gestures that can be used are limited to touch or mid-air ones.

To the best of the authors' knowledge, no study was performed to analyze the impact of different gestures associated with the same TUI, and how this setup can affect the HRI in a PRG. To cope with this unsolved point, the purpose of this study is to investigate the effect of different manipulation

This work has been partially supported by the VR@POLITO initiative.

Fig. 1. Anki Cozmo and bundled cube-shaped TUIs.

interactions associated with the same TUI in a MR-based PRG.

II. Materials and Methods

In an attempt to make the findings derived from this study directly applicable to the consumer level, it was decided to use for the investigation a COTS toy robot. More specifically the *Anki Cosmo*[1] robot was selected among others, due to its popularity, because it has a strong emotional connotation [14] and, more importantly, because it is bundled with a set of interactive physical props shaped as cubes (Fig. 1) that can be easily exploited as TUIs. The cubes can also give feedback in the form of colors. Further technical details about Cozmo will be given in Section II-B.

A. Game Design and Implemented Variants

As said, the aim of this work is to investigate whether different tangible interactions methods can influence the game experience in a PRG, and whether the user's perception about the robot is influenced as well. To this aim, a new PRG named *TangiPong* was designed and implemented.

The game was inspired by two classic games, *Pong* and *Simon Says*. As in the classic *Pong* game, two players control a paddle (one each) by moving it across one side of a rectangular play area in the attempt to hit a ball and score a point. Points are earned when one player fails to return the ball to the other. In *TangiPong*, the paddles are the game "data" associated with a tangible element (a Cozmo's cube), and each player can control its own paddle by physically moving the TUI on the play surface. The play surface was arranged as a tabletop area where the virtual elements are projected from above. The virtual projected elements are: the ball, the playground boundaries, the scores, and some visual feedback (e.g., ball explosion, etc.). It is worth noting that the virtual counterpart of paddles is not projected, both to reinforce the perception of the entanglement between data and props and to force the players to hit the ball by keeping the TUI close to the play surface.

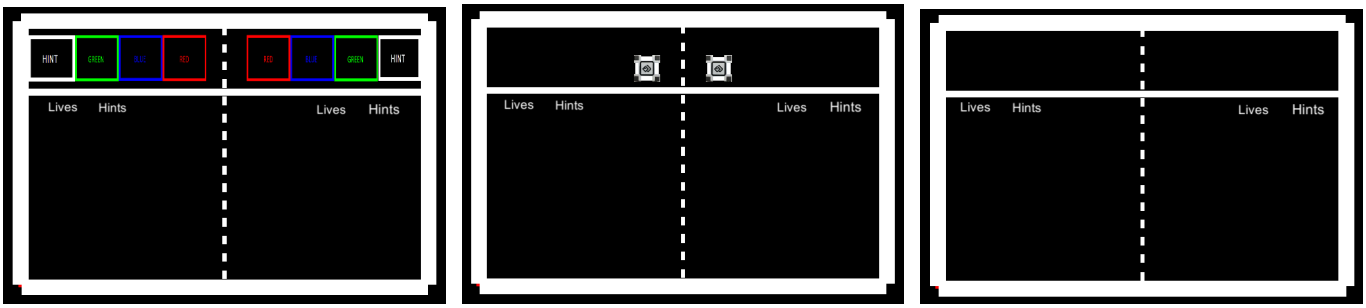
The base mechanics were fused with the ones of *Simon Says*. In particular, each time a player wants to hit the ball, it has to perform the task by previously setting a specific color to its cube (Action 1). The color to set is part of a sequence that is defined for both the players before each round starts. Available colors are red, blue and green. The sequence will be consumed by both players in alternation. For instance, if the given sequence is Red-Blue-..., Player A should hit the ball setting the cube to Red, Player B to Blue, and so forth. At the beginning of each round, the ball has random chances to go to either Player A or Player B, thus forcing both players to remember the entire sequence of colors and not just the odd (even) ones. A round is over when a player scores a point by sending the ball on the other player end zone (the other player loses a life), when a player hits the ball with the wrong color (that player loses a life), or when both the players consumed correctly all the ball exchanges using the correct sequence of colors. In the first two cases, the round is repeated with the same sequence; in the latter case, a new round is started extending the sequence by two additional colors. The game ends when one of the two players loses all the three lives available. The robot fulfills the role of referee similarly to [26]. Hence, it is responsible for providing main feedback to players regarding the reasons of the round conclusion [14], and to inform them about the colors sequence. The sequence is provided by the robot by voice synthesis: no additional visual feedback is provided thus catalyzing the attention of the player on the robot without distractions. The robot usually stays at the border of the play surface, in the middle of the longest side of the rectangular area. If a player is not able to recall the correct current color, it has the faculty to ask the robot for help by interacting with the TUI (Action 2): the robot will inform the player about the correct color to set. Each player can ask for a maximum of four hints, which can be spent without constraints (even all in the same round).

This game behavior was then declined in three variants, illustrated in Fig. 2. Video footage of the three variants are available for download[2]. Variants differs in the specific set of gestures associated to Action 1 and Action 2. In addition, the robot hint provisioning is declined coherently with the associated gesture set as reported below.

- *Place*. Into an area external to the playground but still inside the play surface, four boxes are provided respectively for the three colors and the hint request. Action 1 consists in placing the TUI in the correct color box. Action 2 consists in placing the TUI in the hint box. The robot will provide the hint by moving aside the correct color box and speaking the color loud.
- *Tap*. Into an area external to the playground but still inside the play surface, two images representing the TUI cubes are provided respectively for the two players. Action 1

[1]Anki Cozmo: https://anki.com/en-us/cozmo.html

[2]Video footage of the game variants: http://tiny.cc/vmxcaz



(a) (b) (c)

Fig. 2. The virtual playground of the game variants implemented in *TangiPong*: a) *Place*, b) *Tap*, and c) *Mid-air*.

consists in one-tapping on the TUI while it is not moving. Action 2 consists in double-tapping on the TUI while it is not moving. The time discriminator between the two actions was set to a threshold of 0.25 seconds. The robot, which is standing between the two projected cubes, will provide the hint by rotating towards the cube that matches the player who asked for the hint, tapping on it by using the lifter and by speaking the color loud. When the lifter hits the ground, the color of the virtual cube changes accordingly.

- *Mid-air*. The player has to perform the two actions in mid-air and then put the TUI back to the ground to use it in its normal paddle mode. The idea behind Action 1 is to simulate a knob rotation: hence, it consists in rotating the cube by $90°$ on an axis parallel to the play surface, indifferently clockwise or anticlockwise, and then back to the initial orientation. Action 2 consists in a shake gesture, that is, the TUI has to be shaken vigorously to trigger the action. The robot will provide the hint by turning towards the player position (or better to say, pointing to the last tracked position of the TUI) and by speaking the color loud.

These sets of gestures belong to the group of so called "pragmatic practical action in haptic direct manipulation" [15], [23], and are the most recurrent among tangible interactions applied in tabletop spatial AR. Nonetheless, each set is disjoint from the others owing to the very specific kind of action. It is worth noting that the defined implementation of color selection in Action 1 is, by design, non-homogeneous. In fact, while in the *Touch* and *Mid-air* variants performing Action 1 entails iterating over colors, in the *Place* variant Action 1 is based on a direct selection of the desired color without any iteration. To balance the intrinsic impairment associated with this choice, the ball speed is set as to allow that each player could change colors a minimum of five times through iteration. This one was considered a good compromise as, differently than the original *Pong*, *TangiPong* was designed more as a memory game than a dexterity and coordination game. Furthermore, the goal of this study is not to evaluate implications about dexterity and effectiveness of a given gestures set w.r.t. the others but, rather, to investigate the impact of such different gestures in the game experience and on the relation with the robot.

B. Technologies

As stated in the early Section II, the robotic platform selected for the experiment was the Anki Cozmo (Fig. 1). The manufacturer provides an official SDK[3] for programming it in Phyton. Cozmo is a non-holonomic robot sized $6 \times 7 \times 11$cm at rest, including two moving parts (wheels excluded). A first moving part, which can be considered as the "head" of Cozmo, has one rotational degree of freedom (DOF), and can rotate by $20°$ downward and $45°$ upward. The head of Cozmo is completed by a "face" implemented through a 2×2cm LED matrix display, which shows a simplified anthropomorphic facial expression using eye-like animations (which can be selected from a pre-defined list using the SDK). Beneath the display, there is a $60°$ wide field of view 640×480 pixels RGB camera (although the image accessible through the SDK is limited to a 320×240 grayscale image). The second moving part is a front lifter (one positional DOF, likewise controllable through the SDK), which is designed to interact with the interactive cubes that are bundled with the robot and can be lifted by 5.5cm using this tool. The three cubes of 5cm edge, are equipped with 4 RGB LEDs arranged on the 4 edges of one face of the cube (the cube color can be changed through the SDK). The cubes are also equipped with a three DOF accelerometer, whose data are accessible through the SDK as well. The cubes can be distinguished by different markers that are used to track them relative to the robot using its front camera. Cozmo is also equipped with WiFi capabilities and a built-in speaker. The SDK is designed using an event-driven approach and is rich in features (for the sake of brevity, in this review only the subset of features that were actually used for the implementation were mentioned).

The complete high-level architecture of the spatial AR-based system exploited in this work is illustrated in Fig. 3. The setup, shown in Fig. 4, included Cozmo, two cubes, a RGB camera, a projector, an Android smartphone and a PC running Windows 10. The projector was mounted near the ceiling in order to project the image on the table from the top. To improve the quality of the projected image, the table was covered using a black cardboard of size 100×65cm, which is also the size of the play surface. In the immediate

[3]Cozmo's SDK: http://cozmosdk.anki.com/docs/index.html

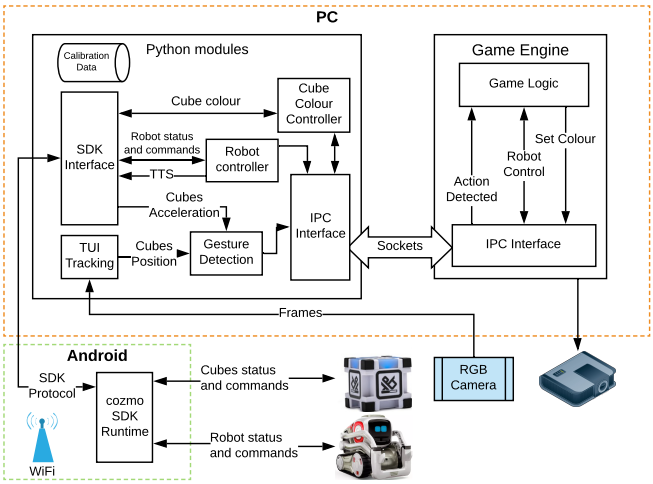


Fig. 3. Architecture of the gaming system.

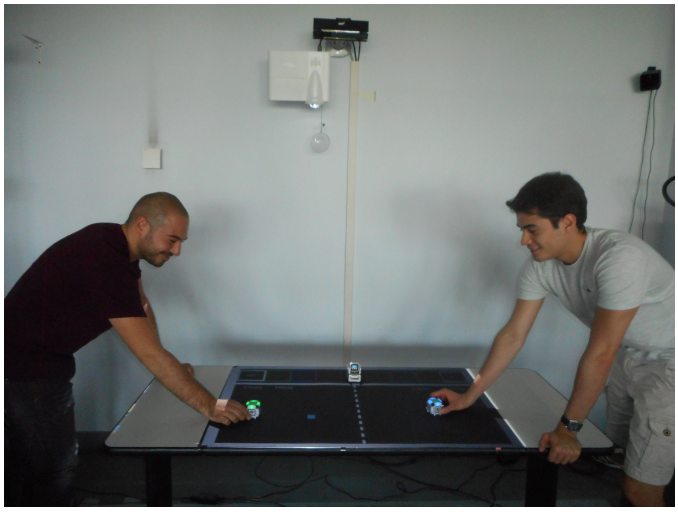


Fig. 4. Setup of the gaming system.

nearby of the projector it was mounted a Microsoft Kinect v2, with the same orientation. For this specific setup, just the 1920×1080 pixels, 30fps RGB camera of the latter device was used to track the position of the cubes in the play area. In fact, the quality of Cozmo's built-in tracking functionality is strictly constrained by the fact that the cube needs to be seen by the robot. Thus, it was preferred to implement an external *outside-in* tracking system for the cubes. On each cube it was attached a green colored 3×3 cm carboard, which was used to easily locate the cube in the play area using well-known computer vision techniques (color range thresholding, opening, shape detection). A calibration phase (performed before the game starts) was required to synchronize the Cozmo's internal coordinate system with the coordinate system used by the external tracking and by the projection, computing the required transformation matrices. Due to the unreliability of the accelerometer data provided by the event-based SDK of the robot, the sensor fusion option was excluded. These data were reliable enough, indeed, to be exploited by the gesture recognition module. Voice feedback was provided (when requested) using the text-to-speech capabilities of the SDK in English language. The game logic and graphics were implemented using the well-known Unity game engine, and were deployed to a Windows application running on the PC. The gesture detection module, the tracking and the robot control logic were instead implemented in another Python application, accessing the functionalities provided by the Cozmo's SDK. The inter-process communication was implemented through sockets. The cubes and Cozmo communicate using a WiFi connection hosted by the smartphone. The SDK, in fact, requires a runtime in execution on an Android (or iOS) device, which has to be connected to the PC running the applications through USB cable. In this way, the smartphone acts as a communication interface between the PC and the robot with the TUI.

III. Results and Discussion

This section presents an discusses the results of a preliminary user study that was carried out by using the devised game to evaluate the influence of the different set of gestures on the user's perception of the game experience and relation with the robot.

The population of the study included 20 volunteers (13 males and 7 females) aged between 20 and 29 y.o. ($\mu = 23.35, \sigma = 3.34$), selected among university students. Volunteers were randomly coupled, and each couple was requested to play all the three variants of the game; the order in which the three variants were played was randomized with the aim to reduce possible learning effects. When the variant was changed, the players were also requested to switch their positions. For each variant, a game was considered as valid if both players had performed at least one time both Action 1 and Action 2. No constraint was given about a minimum round to reach (or, in other words, a minimum length of the sequence of colors to remember), which was anyway quite high ($\mu = 6.3, \sigma = 1.4$)

Before the experience, all the players were asked to respond to a pre-test questionnaire designed to investigate their previous knowledge and expertise with technologies related to those used in the game. According to information collected, volunteers were particularly used to play video-games; in particular, 63.2% of them said to play video-games regularly (every day, once a month or once a week), whereas 26.3% just a few times. Concerning the use of toy robots, 84.2% of volunteers said have never used these technologies and 63.2% never used other kinds of service robots. Natural interfaces, e.g., for hand and body gesture-based control were more familiar, with 52.6% of the volunteers who had used them few times, 15.8% once a month or once a week and 5.3% every day, whereas 26.3% never used such interfaces.

After the experience, players were asked to respond to a post-test questionnaire by expressing their agreement with

TABLE I
SELECTION OF THE MOST RELEVANT RESULTS FROM THE QUESTIONNAIRE.
VARIANTS ARE REPORTED ABBREVIATED AS P (PLACE), T(TAP) AND MA (MID-AIR)

Statement	μ			σ			$D_{m,n} < D_{critic}$ (K-S)			Concordance (P)		
	P	T	MA	P	T	MA	P-T	P-MA	T-MA	P-T	P-MA	T-MA
SUS score	84.13	81.38	78.75	9.40	10.35	12.46	yes	yes	yes	0.06	-0.34	0.32
I found the various functions in this gaming system were well integrated	3.20	3.30	3.10	0.68	0.56	0.62	yes	yes	yes	0.77	0.42	0.63
It was easy to get the desired colour	3.70	3.20	2.90	0.56	0.75	1.04	yes	yes	yes	0.14	-0.40	0.47
It was easy to request the suggestion from the robot	3.85	3.10	2.85	0.48	0.94	1.01	no	no	yes	0.03	-0.15	0.33
It was easy to confuse the action of changing colour with the action of requesting the suggestion	0.70	1.55	0.90	0.64	0.92	0.83	no	yes	yes	0.20	0.41	0.46
The interaction method is pleasant to use	3.60	3.30	3.25	0.58	0.71	0.77	yes	yes	yes	-0.07	-0.45	0.59
I was able to hit the ball as expected	3.85	3.75	3.80	0.48	0.77	0.60	yes	yes	yes	0.85	0.94	0.98
I perceived the robot as an intelligent part of the gaming system	3.65	3.55	3.45	0.57	0.59	0.67	yes	yes	yes	0.87	0.67	0.77
I perceived the robot as an enemy	0.25	0.25	0.25	0.43	0.43	0.43	yes	yes	yes	1.00	1.00	1.00
In general. I felt the robot presence was crucial for the game	2.95	2.90	2.90	1.07	1.14	1.14	yes	yes	yes	0.98	0.98	1.00
I felt robot has a personality	2.80	2.80	2.80	0.87	0.87	0.87	yes	yes	yes	1.00	1.00	1.00
I felt like the robot was intentionally reacting to my actions	2.15	2.10	2.10	0.96	0.94	0.94	yes	yes	yes	0.97	0.97	1.00
I clearly understood the suggestion provided by the robot when needed	3.80	3.75	3.60	0.51	0.43	0.66	yes	yes	yes	0.45	0.35	0.87
I think the way the robot suggest me the colour was coherent with the game version	3.95	3.95	3.80	0.22	0.22	0.40	yes	yes	yes	-0.05	-0.10	0.22
I liked the way the robot was moving	3.65	3.70	3.65	0.57	0.56	0.57	yes	yes	yes	0.93	0.85	0.93
I clearly understood what the robot was saying	3.35	3.30	3.25	0.79	0.78	0.77	yes	yes	yes	0.96	0.93	0.96
I wold have preferred a different way to provisioning the suggestion WITH the robot	0.90	0.90	1.10	1.18	1.09	0.94	yes	yes	yes	0.96	0.77	0.79
I wold have preferred a different way to provisioning the suggestion WITHOUT (instead of) the robot	0.75	0.65	0.75	0.94	0.79	0.94	yes	yes	yes	0.89	0.77	0.89
I would have preferred a more direct interaction with the robot (eg. Touching. hitting.)	1.75	1.75	2.00	1.13	1.13	1.18	yes	yes	yes	1.00	0.86	0.86
I felt challenged by the game	3.05	2.95	2.95	0.74	0.76	0.74	yes	yes	yes	0.29	0.64	0.61
I enjoyed the game	3.35	3.30	3.20	0.91	0.90	0.93	yes	yes	yes	0.85	0.69	0.89
What the other(s) did affected what I did	2.05	2.05	2.05	1.07	1.07	1.07	yes	yes	yes	0.96	0.96	1.00
I paid close attention to the other(s)	3.00	3.10	3.00	0.55	0.54	0.55	yes	yes	yes	0.85	1.00	0.85
What I did affected what the other(s) did	2.40	2.50	2.40	1.02	0.97	1.02	yes	yes	yes	0.96	1.00	0.96
Overall I liked the game	3.70	3.55	3.35	0.46	0.50	0.79	yes	yes	yes	0.29	0.29	0.65

several statements on a 0 to 4 scale (from strongly disagree to strongly agree). The post-test questionnaire consisted of four sections. The first section included all the questions of the *System Usability Scale* tool [27]. The second section, named *Tangible interaction*, explored the way players perceived the TUI to interact with the gaming system (e.g., if the interaction method was pleasant to use, if it was easy to set the desired color, etc.) The third section, named *Robot perception*, investigates the players' perception about the role of the robot and its interaction with them. In this section, is was explored how the different ways for providing suggestions was perceived. The fourth section, named *Game experience*, explored the way players perceived the game experience (e.g., if they felt bored or challenged, what was their sentiment about the interactivity of the game considering the robotic element and the projected playground) and the relationship with the opponent. Questions were adapted from the CEQ questionnaire [28]. Lastly, players were asked to express and motivate their preference for the three game variants, by providing open feedback about positive and negative aspects for each of them. The questionnaire is available for download[4].

Collected feedback was analyzed by means of Kolmogorov-Smirnov tests (K-S) on pairwise variants. In addition, in order to evaluate whether there was any concordance over the testers among a specific question, Pearson's correlation analysis (P) was used. A selection of the most relevant results is provided in Table I.

The general finding of the study is that the three gestures sets did not have a significantly different impact on the game experience and on robot perception. Concerning the tangible interaction, users found a significant difference in the easiness of confusing Action 1 and Action 2 in the *Tap* variant compared other two variants (*Tap* $= 1.55$, *Place* $= 0.7$, *Mid-air* $= 0.9$). While still fairly acceptable, this fact implicates that, in case of multiple actions, tapping should be discarded in favor of other sets of gestures, or that coherency within the set should be sacrificed in favor of naturalness (e.g., substituting double-tap with turning). Also, there is a significant difference in the *Place* variant w.r.t. the other two variant about whether it was easy to perform Action 2 (*Place* $= 3.85$, *Tap* $= 3.1$, *Mid-air* $= 2.85$). It is worth noting that in both cases the concordance among user was scarce ($P < 0.5$ for each pair).

The SUS score is very high in all the three variants (*Place* $= 84.1$, *Tap* $= 81.4$, *Mid-Air* $= 78.8$), and there is no significant difference among them, thus possible preference for one of the variants analyzed in a dedicated section of the questionnaire should not be attributed to a gap in the usability or in the implementation quality.

There was very high concordance ($P > 0.8$) among players about the fact that they were able to hit the ball as they expected ($\mu = 3.8$), about the central role of the robot in the game ($\mu = 2.9$), about the perception of the robot as a lifelike being ($\mu = 2.55$), and about its personality ($\mu = 2.8$).

Furthermore, users found that the suggestion given by the robot was pretty clear in all the variants, as well as that is was coherent with the gestures set.

About the game experience, the game was judged as satisfactory at stimulating the interest of players and their competitive behaviors, and no significant difference was spotted among the three variants with high concordance. There is no significant difference also in the overall game likability of the three variants, though with low concordance. This is reflected also by the answers about the overall preference ("Which variant did you prefer?"), as 55% of the users preferred the *Place* variant over the *Tap* (20%) and the *Mid-air* (25%) ones.

Lastly, based on the open feedback, players complained about the (low) speed of the ball, appreciated the (referee) role of the robot, but would have preferred the introduction of

[4]Questionnaire: http://tiny.cc/vwxcaz

a more direct physical interaction with it (e.g. "by giving it high five for the winning feedback").

IV. Conclusions and Future Work

In this paper, it was presented a study designed to investigate the influence that a diverse set of gestures entangled to an invariant TUI can have on the user's perception of the game experience and on relation with a robot. A new MR-based PRG leveraging tabletop projection was designed and implemented for that specific purpose, by involving the use of TUIs.

From a preliminary user study conducted on three variants of the game, it emerged that there is no significant difference in using one gesture set compared to the others, and that the selection of one set does not impact significantly on the game experience and on the perceived role of the robot. From the TUI action per se, it emerged that the usage of interaction involving tap should be considered as less natural when multiple actions are required (e.g. tap, double-tap) w.r.t. other sets of gestures.

Future work will be focused on extending the experimental evaluation to a larger number of subjects, as well as on expanding the number of variants to compare. In addition, taking also into account collected feedback, it will be considered to implement more direct mechanics for interacting with the robot, and to give also haptic feedback when the TUIs interacts with virtual elements.

References

[1] K. Baraka, P. Alves-Oliveira, and T. Ribeiro, "An extended framework for characterizing social robots," *arXiv preprint arXiv:1907.09873*, 2019.

[2] M. A. Goodrich, A. C. Schultz *et al.*, "Human–robot interaction: a survey," *Foundations and Trends® in Human–Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2008.

[3] D. Martinoia, D. Calandriello, and A. Bonarini, "Physically interactive robogames: Definition and design guidelines," *Robotics and Autonomous Systems*, vol. 61, no. 8, pp. 739–748, 2013.

[4] M. L. Lupetti, G. Piumatti, and F. Rossetto, "Phygital play hri in a new gaming scenario," in *2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN)*. IEEE, 2015, pp. 17–21.

[5] G. Piumatti, A. Sanna, M. Gaspardone, and F. Lamberti, "Spatial augmented reality meets robots: Human-machine interaction in cloud-based projected gaming environments," in *2017 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2017, pp. 176–179.

[6] D. Calife, J. L. Bernardes Jr, and R. Tori, "Robot arena: An augmented reality platform for game development," *Computers in Entertainment (CIE)*, vol. 7, no. 1, p. 11, 2009.

[7] B. Lahey, W. Burleson, C. N. Jensen, N. Freed, and P. Lu, "Integrating video games and robotic play in physical environments," in *Proceedings of the 2008 ACM SIGGRAPH symposium on Video games*. ACM, 2008, pp. 107–114.

[8] D. Robert and C. Breazeal, "Blended reality characters," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 2012, pp. 359–366.

[9] K. Sakai, Y. Hiroi, and A. Ito, "Playing with a robot: Realization of" red light, green light" using a laser range finder," in *2015 Third International Conference on Robot, Vision and Signal Processing (RVSP)*. IEEE, 2015, pp. 1–4.

[10] D. Robert, R. Wistorrt, J. Gray, and C. Breazeal, "Exploring mixed reality robot gaming," in *Proceedings of the fifth international conference on tangible, embedded, and embodied interaction*. ACM, 2011, pp. 125–128.

[11] H. Costa, P. Cebola, T. Cunha, and A. Sousa, "A mixed reality game using 3pi robotspitanks," in *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 2015, pp. 1–6.

[12] F. Lamberti, D. Calandra, F. Bazzano, F. G. Prattico, and D. M. Destefanis, "Robotquest: A robotic game based on projected mixed reality and proximity interaction," in *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 2018, pp. 1–9.

[13] F. Lamberti, A. Cannavò, and P. Pirone, "Designing interactive robotic games based on mixed reality technology," in *2019 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2019, pp. 1–4.

[14] F. Lamberti, F. G. Pratticò, D. Calandra, G. Piumatti, F. Bazzano, and T. R. Villani, "Robotic gaming and user interaction: Impact of autonomous behaviors and emotional features," in *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 2018, pp. 1–9.

[15] E. Hornecker and J. Buur, "Getting a grip on tangible interaction: a framework on physical space and social interaction," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 437–446.

[16] L. E. Holmquist, A. Schmidt, and B. Ullmer, "Tangible interfaces in perspective," 2004.

[17] G. W. Fitzmaurice and W. Buxton, *Graspable user interfaces*. University of Toronto, 1997.

[18] M. Sugimoto, T. Fujita, H. Mi, and A. Krzywinski, "Robotable2: a novel programming environment using physical robots on a tabletop platform," in *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*. ACM, 2011, p. 10.

[19] T. Sapounidis and S. Demetriadis, "Tangible versus graphical user interfaces for robot programming: exploring cross-age children's preferences," *Personal and ubiquitous computing*, vol. 17, no. 8, pp. 1775–1786, 2013.

[20] A. Krzywinski, H. Mi, W. Chen, and M. Sugimoto, "Robotable: a tabletop framework for tangible interaction with robots in a mixed reality," in *Proceedings of the International Conference on Advances in Computer Enterntainment Technology*. ACM, 2009, pp. 107–114.

[21] K. P. Fishkin, "A taxonomy for and analysis of tangible interfaces," *Personal and Ubiquitous computing*, vol. 8, no. 5, pp. 347–358, 2004.

[22] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The grasp taxonomy of human grasp types," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 66–77, 2015.

[23] E. van den Hoven and A. Mazalek, "Grasping gestures: Gesturing with physical artifacts," *AI EDAM*, vol. 25, no. 3, pp. 255–271, 2011.

[24] C. Guo and E. Sharlin, "Exploring the use of tangible user interfaces for human-robot interaction: a comparative study," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2008, pp. 121–130.

[25] J. Ruiz, Y. Li, and E. Lank, "User-defined motion gestures for mobile interaction," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 197–206.

[26] G. Piumatti, F. G. Prattico, G. Paravati, and F. Lamberti, "Enabling autonomous navigation in a commercial off-the-shelf toy robot for robotic gaming," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2018, pp. 1–6.

[27] J. Brooke *et al.*, "Sus-a quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[28] W. IJsselsteijn, Y. De Kort, and K. Poels, "The game experience questionnaire," *Eindhoven: Technische Universiteit Eindhoven*, 2013.