Mixed Reality Tabletop Gameplay: Social Interaction with a Virtual Human Capable of Physical Influence

Myungho Lee, *Member, IEEE,* Nahal Norouzi, *Student Member, IEEE,* Gerd Bruder, *Member, IEEE,* Pamela J. Wisniewski, *Member, IEEE,* and Gregory F. Welch, *Senior Member, IEEE*

Abstract—In this paper, we investigate the effects of the physical influence of a virtual human (VH) in the context of face-to-face interaction in a mixed reality environment. In Experiment 1, participants played a tabletop game with a VH, in which each player takes a turn and moves their own token along the designated spots on the shared table. We compared two conditions as follows: the VH in the *virtual* condition moves a virtual token that can only be seen through augmented reality (AR) glasses, while the VH in the *physical* condition moves a physical token as the participants do; therefore the VH's token can be seen even in the periphery of the AR glasses. For the *physical* condition, we designed an actuator system underneath the table. The actuator moves a magnet under the table which then moves the VH's physical token over the surface of the table. Our results indicate that participants felt higher co-presence with the VH in the *physical* condition, and participants assessed the VH as a more physical entity compared to the VH in the *virtual* condition. We further observed transference effects when participants attributed the VH's ability to move physical objects to other elements in the real world. Also, the VH's physical influence improved participants' overall experience with the VH. In Experiment 2, we further looked into the question how the physical-virtual latency in movements affected the perceived plausibility of the VH's interaction with the real world. Our results indicate that a slight temporal difference between the physical token reacting to the virtual hand's movement increased the perceived realism and causality of the mixed reality interaction. We discuss potential explanations for the findings and implications for future shared mixed reality tabletop setups.

Index Terms—Augmented reality, virtual humans, physical-virtual interaction, latency.

1 Introduction

Augmented reality (AR) technologies have seen major advances over the last years with developments such as the Microsoft HoloLens and generally less expensive and more usable displays, sensors, and user interfaces [1]. While not there yet, it seems reasonable to assume that AR displays will become a common sight for home cinema, gaming, and related experiences over the next decade. In particular in peoples' homes, AR technologies can have a strong impact on how we interact with each other, e.g., using AR telepresence [2], and with virtual humans (VHs), such as embodied forms of intelligent virtual agents [3]. The advent of voice-controlled agents over the last years and their embodied AR counterparts have shown the potential of such agents to act as social entities in our daily life [4]. Such VHs can take on a plethora of roles that are typically taken by real humans in our daily lives, such as assistants, companions, supporters, or adversaries, e.g., when playing a tabletop game alone or in a group at home.

However, when interacting with a VH that is presented via optical see-through glasses such as the HoloLens, the challenge remains that the virtual content is not able to exert a direct influence over the physical entities in the room. This can have a negative effect on users' sense of *co-presence*, which is defined as "the degree to which one believes that he or she is in the presence of, and dynamically interacting with, other veritable human beings" [5], [6]. Harms and Biocca described co-presence as one of several dimensions that make up *social presence*, i.e., one's sense of being socially connected with the other [7].

In this paper, we present a technical approach to realize physical-virtual interactivity in AR in the scope of a tabletop environment, and we present an example application and user study designed around a tabletop gaming experience between a real and a virtual human. The study involved two conditions in which the VH either exerted influence over *physical* or *virtual* tokens on the tabletop surface. With subjective and behavioral measures, we show benefits of the physical condition on the participants' sense of co-presence as well as their sense that the VH is a physical entity.

This paper is structured as follows. Section 2 presents related work in the scope of VHs and physical-virtual interactivity. Section 3 describes the apparatus and tabletop setup that we developed to give the virtual content control over the movement of physical objects on a tabletop surface. Section 4 describes the human-subject study in which we investigate the benefits and drawbacks of such an influence. Section 5 presents a follow-up study study in which we investigate the effects of physical-virtual latency on the perceived plausibility of such interactions. Section 6 concludes the paper and discusses future work.

- *M. Lee is with the Korea Institute of Science and Technology, Seoul, South Korea. E-mail: mlee@kist.re.kr*
- *N. Norouzi, G. Bruder, P. J. Wisniewski, and G. F. Welch are with the University of Central Florida, 3100 Technology Parkway, Orlando, FL 32826–3281, United States.*
  E-mail: see https://sreal.ucf.edu/people

Manuscript received March 15, 2019.

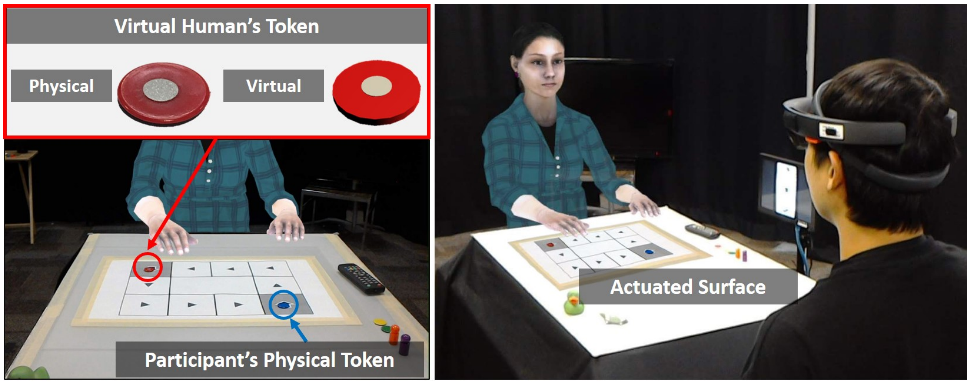


Fig. 1. Illustration of the augmented reality game mechanics with virtual or physical game tokens on the left. The image on the right side shows the tabletop gaming surface with the magnetic actuator system underneath, which gives the illusion of the virtual human being able to touch and move physical objects over the surface.

2 Related Work

In this section we resume related work on VHs in AR and their interaction with physical entities in the real space.

2.1 Virtual Humans in AR

The term *virtual human* generally refers to human-like computer graphics manifestations. They can appear in a virtual environment or can share a physical space with real humans. Traditionally virtual humans are referred to as *avatars* or *agents* depending on the entity controlling them, where avatars are controlled by humans, while agents are controlled by computer programs [8]. Various application fields employ and draw benefits from VHs (see [9]). For example, Hoque et al. [10] developed a system for users to train their social skills, e.g., job interview skills, with VHs that could give personalized feedback. Because of the resemblance of VHs' appearance and shape with humans, people naturally distinguish them from non-human objects and often treat them in a similar way as real humans [11]–[13]. The phenomenon that people treat VHs as if they were real humans is often leveraged in training simulations, where they assume the roles of instructors or training partners that may not always be available.

Social presence and co-presence are commonly used constructs to measure users' perception of VHs. They are generalizable factors among many other simulation-dependent factors in assessing the effectiveness of training simulations that employ VHs. While many interpretations of the terms social presence and co-presence have been proposed (see [14]), Goffman et al. [15] indicated that co-presence exists when people feel that they are able to perceive others and that others are able to perceive them. Harm and Biocca [7] defined social presence as "one's sense of being socially connected with the other" and "one's sense of the other person's presence."

Researchers have investigated traits of VHs, e.g., appearance, shape, realistic gestures, to increase users' sense of social and co-presence. However, a relatively small amount of research has attempted to bring realistic three-dimensional VHs in users' physical environment in AR [16], compared to the majority of research performed in Virtual Reality (VR). Increasing convergence of AR and Robotics in different areas such as using AR as a social interface for a robot [17], robot path planning [18], or implementing a VH's autonomous behavior such as eye and head motion [19] through the advances of the same topic in the field of robotics [20], [21], can provide a turning point in AR research. Meanwhile, efforts to make a social robot, e.g., for a human companion, has been steadily made in the robotics community [22], but they faced Uncanny Valley related challenges due to the complexity of representing realistic human facial expression as well as subtle body gestures [23]. Convergence of AR and robotics, i.e., the realistic 3D graphics of AR and the physical presence of robots, in this regard, might be mutually beneficial for both VHs in AR and social robots [24]. When VHs are brought into users' real space, two main approaches exist: (i) They can be partially or entirely projected onto physical objects that look like a human body, or (ii) they can be overlaid onto a user's view using AR technology. For example, Kotranza et al. [25] proposed a mannequin-based embodied agent, a virtual patient, that supports touch interaction with medical trainees. Similarly, Lincoln et al. [26] prototyped a robot-based embodied virtual human. They projected a human face onto an actuated robotic head which could convey non-verbal social behavior, such as gaze direction, as well as verbal communication. Obaid et al. [27] used video see-through AR glasses to augment the VH in a user's view in their study evaluating the relationship between the users' physiological responses and VHs' cultural behaviors.

However, there are perceptual issues one should consider when using AR glasses to overlay VHs in the users' view (see [28]). For instance, Lee et al. [29] showed that the small augmented field of view of the current-state optical see-through AR glasses can affect users' proxemic behavior in the presence of VHs. Also, Kim et al. [30] indicated that VHs' conflicting physical behavior with real objects, e.g., passing through them, could reduce users' sense of co-presence with the VH.

2.2 Physical-Virtual Interactivity

Bridging the gap between the physical world and virtual worlds has been of increasing interest in recent years. For instance, Sra et al. [31] introduced a method to create a walkable area in a virtual environment that is based on the space in the real world. Similarly, Simeone et al. [32] proposed a *substitutional reality* where the physical world is substituted with virtual counterparts, and showed a relation between the level of mismatch and the user experience in such an environment. Regarding the opposite direction, from virtual to real, researchers have proposed methods utilizing mobile robots and actuators. He et al. [33] demonstrated three different mapping mechanisms between physical and virtual objects in such scenarios. Kasahara et al. [34] proposed "exTouch", a touchscreen-based interaction method, to allow users to manipulate actuated physical object through AR. Joshua et al. [35] used networked actuators to bring virtual events into the physical world in their *cross reality* implementation.

Unlike VR, however, in augmented/mixed reality, virtual content is overlaid onto or mixed with the real world, creating a unified world. In such cases, the means by which virtual entities interact with the physical environment can affect users' perception. For example, Kim et al. [30] demonstrated that users rated the sense of social presence higher with a VH that exhibited awareness of the physical space, compared to one that did not in AR. This finding is comparable to the results of Bailenson et al. [11], in which a VH that exhibited awareness of the user in an immersive virtual environment received higher social presence and induced more realistic gaze and proxemic behavior with the participant.

Similarly, users had higher co-presence with a VH that could affect their physical space. Lee et al. [36] showed that participants rated co-presence higher with a VH when it could affect their physical space through a shared physical-virtual table in a mixed reality environment. They used an actuated wobbly table to establish such physical-virtual interactivity. Later, Lee et al. [29] also showed that subtle tactile vibrations of a VH's footsteps could induce higher co-presence with the VH in AR.

We are entering an era where VHs can be given more and more control over physical objects at our homes and in public spaces. With the Internet of Things (IoT), common devices in our daily lives are connected to computer systems that enable them to be accessed by voice-controlled agents, such as Amazon Alexa, providing an intuitive and natural interface to interact with them [4]. For instance, Kim et al. [3] investigated IoT devices as a VH's physical influence channel and compared the effects of embodied voice-controlled agents and their behavior on the user experience as well as social presence. They found that exhibiting plausible behavior, e.g., walking over to an IoT lamp and pretending to touch a light switch to turn it on, similar to what real humans would do, induced significantly higher social presence with the agent than voice-only interaction.

In addition to those IoT devices, some tangible interfaces seem promising candidates for realizing physical-virtual interactivity for VHs. For example, Follmer et al. [37] developed a shape-changing surface with a grid of linear actuators and demonstrated various interaction techniques using the surface. Leithinger et al. [38] later used the shape-changing surface to allow two remote users to interact physically through the surface. The actuated surface in this paper is further inspired by the work by Lee et al. [39]. Though they did not consider AR or VHs, they presented an approach based on an electromagnet with a three-dimensional actuated stage to levitate a ball-shape permanent magnet in mid-air.

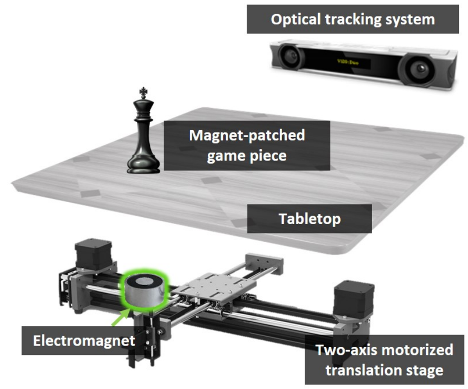


Fig. 2. Apparatus: Tracked magnetically actuated game pieces on a tabletop surface realized through a motorized translation stage hidden from view underneath the surface.

3 Apparatus

This section describes the tabletop setup with the magnetic actuator system underneath the surface that we developed for use with a virtual human presented in AR (see Figure 1).

3.1 Magnetic Actuator Surface

We designed an apparatus that can extend the ability of VHs in AR to move physical objects on a surface (see Figure 2).

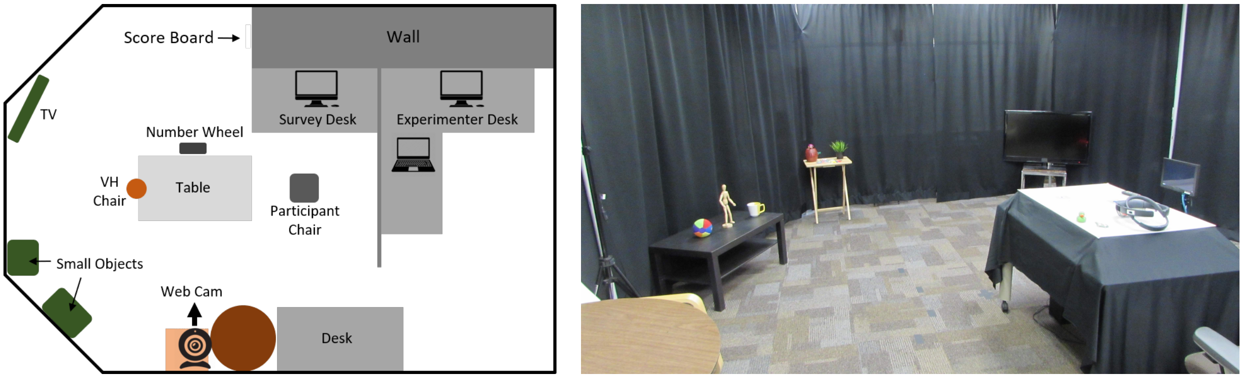
The apparatus comprises the four main components:

- A magnet that can attract magnet- or metal-patched physical objects on the surface of the table.
- A two-axis motorized *translation stage* that can move the magnet parallel to the surface of the tabletop.
- A tracking system that tracks the positions of physical objects on the table and sends the data to AR glasses to register virtual content accordingly.
- A tabletop that covers the translation stage and hides it from the user's view.

We used an EleksDraw Computer Numerical Control (CNC) machine for the two-axis motorized translation stage and mounted a magnet to the mobile part of the CNC machine at the tip where usually a drill or laser is attached. The working range of the translation stage is 280 mm $\times$ 200 mm, and the maximum speed is 83 mm/s. We used an ease in/out velocity curve for a natural movement of the token; the average speed of the token was 50 mm/s. We compared different electromagnets and permanent magnets, and we decided to use a robust permanent magnet (a neodymium magnet) for the study presented in this paper due to trade-offs between its magnetic force, the weight of the physical object on the surface, and the thickness of the surface.

We used an OptiTrack Duo optical tracking system to track the position of the physical objects on the surface. We mounted the cameras on the ceiling of the experimental space, looking down at the tabletop surface. As the OptiTrack system requires retroreflective infrared (IR) markers to track the position of objects, we attached small markers to the corners of the tabletop and to the game tokens.

We decided to use a Microsoft HoloLens, an optical see-through head-mounted display (HMD), and the Unity 2017.2.1f1 graphics engine for rendering virtual content and presenting it to the user.



(a) (b)

Fig. 3. Experimental setup: (a) Illustration of the experimental space with the tabletop setup and other furniture and equipment, and (b) photo of the room with the tabletop gaming setup.

3.2 Tabletop Gaming Setup

Our AR setup is inspired by a two-player tabletop gaming setup, in which a real human and a virtual human sit on opposite sides of a table and take turns to move their tokens over the tabletop surface with the intention to win a rudimentary board game.

We mounted the magnetic actuator system on a 70 cm × 114 cm table surface in our experimental space (see Figure 1). On the actuated surface, we placed a board game map (24 cm × 32 cm) that contained ten designated fields for game tokens to be placed. The fields were arranged in a rectangle around the board. The size of each field was 8 cm × 8 cm. Each player started on a different field. We marked the starting positions for the VH and participant as well as the direction to move the tokens on the map. The starting positions of the tokens were located on the rightmost side of the row near each player, on opposite ends of the board. The tokens had to be moved in counterclockwise direction around the board. The player who completed a round and reached the starting position with their token first was declared the winner of the round.

A small monitor was placed next to the table to indicate whose turn it is (i.e., either the participant's or VH's) and the number of fields to move the token. We decided not to use physical dice for the game in our setup for the purpose of the experiment due to the fact that this would introduce an element of randomness to the study. Instead, we decided to use a computer-controlled virtual number wheel (similar to that of a slot machine), which was rendered in Unity and presented on the monitor. The numbers presented by the number wheel appeared random to the participants but they were predetermined and counterbalanced in our study.

For the VH to move a physical token on the tabletop surface, we attached a thin magnet (diameter: 20 mm) to the bottom of the token (diameter: 22 mm) and an unobtrusive flat IR-reflective marker on top (see markers shown in Figure 1). The tracked marker positions were streamed to the HoloLens. When it was the VH's turn, the VH first placed her right hand on the tracked position of the token, then the motorized translation stage underneath the table moved the magnet from the current position to the target position, which resulted in the token moving over the tabletop surface. Due to the smooth surface of the board game, the token slided over the table without any noticeable friction. The VH's right-hand position was updated in real time based on the tracked marker position, and inverse kinematics was applied for the upper body posture while the token was moving. Latency between the physical and virtual movements was in average 140 ms.

For the virtual human player, we used an ethnically ambiguous female character that could perform predetermined gestures and had multiple dialogue options for the game scenario. The character was modeled and rigged in Autodesk Maya and animated in the Unity graphics engine. For the VH's speech we hired a female actor to record audio for the dialogues. The gestures and dialogues were linked to the stage of the game. Since the progression of the game was predetermined, the actions could be advanced automatically without noticeable delays with minimal help by a human controller using a GUI-based desktop application. For example, while the number wheel was rotating on the small monitor, the VH moved her head and eyes to look at the wheel and responded appropriately to the result such as by saying, "Oh! I got a three." or "Yes! I am almost done."

4 Experiment 1

In this section we describe the experiment that we conducted to investigate differences between purely *virtual* and *physical-virtual* interactions between a VH and other objects.

4.1 Participants

34 participants (11 female, 23 male, age 18–36, average 23.6) volunteered for this paid study through an advertisement posted at the local university. 11 participants had normal vision and 23 participants had corrected-to-normal vision, either using contact lenses (8 participants) or glasses (15 participants). Participants used a 7-point scale (1=no expertise to 7=expert) to rate their level of familiarity with VR (average 4.5), AR (average 3.79), VHs (average 2.5), and tabletop games (average 5.9). 27 participants ranked their level of computer expertise as proficient or expert.

4.2 Material

In this experiment, we used the physical setup, virtual human, and Unity rendering environment described in Section 3. Verbal interaction between the participant and the VH is performed while wearing headphones of type Sony

MDR-ZX110NC. Ambient noise (a sound recorded from a café) was played via the headphones to render the humming background noise of about 40–46 dB caused by the current realization of the apparatus imperceptible, assuming that it could have an effect on the results.

4.3 Method

We used a within-subjects design. Participants experienced both conditions in randomized order. The conditions were:

C_V The VH moved a *virtual* token.
C_P The VH moved a *physical* token.

Participants moved their physical token by themselves in both conditions.

4.3.1 Procedure

Before the experiment, the experimenter asked participants to read an informed consent form and briefed them on the study and protocol. Once participants gave the informed consent, they donned the HoloLens and went through the procedure of the HoloLens' interpupillary (IPD) calibration app. The experimenter helped participants to correctly adjust the Hololens on their head. Participants filled out a pre-questionnaire that contained demographics questions as well as questions about their prior experience with AR, VR, VHs, and tabletop gaming.

The experimenter then left the experimental room, and the participants started their first game. We used the tabletop gaming scenario described in Section 3.2. Participants played the game with the VH once for each of the two conditions in randomized order. We designed two sequences, depending on the sequence chosen for each game, the numbers in that sequence were displayed sequentially on the small monitor next to the table. The VH started the game both times and according to the number, players and the VH took their turns one after another. Each turn, they advanced their token by the number of steps displayed each time on the screen. In order to be comparable between both conditions, we decided that the VH should win both games.

When the game ended, the experimenter re-entered the room and helped them take off the HoloLens and asked them to fill out a post-questionnaire. Participants then repeated the same procedure for the second condition.

Upon completion of both games, participants were asked to fill out a comparative questionnaire with also contained open questions. Participants then received a monetary compensation for their participation.

4.3.2 Subjective Measures

We measured the following items at the end of each game.

Co-Presence: We used Basdogan's Co-Presence Questionnaire (CPQ) [40] to measure the level of "togetherness," being present together, experienced by the participants while playing the game with the VH.

Perceived Physicality: For this measure, we prepared photos of 15 objects with small, medium, and large sizes. 11 objects were placed inside the experimental area, of which 4 objects were placed on the game table (see Figure 3); remaining 4 object photos were obtained from the internet. Each photo of an object appeared with a sentence, "She can move the object below," and participants were asked to rate their reaction to the sentence using a 7-point Likert scale (1=strongly disagree, 7=strongly agree). For the analysis we used two grouping criteria: size (small, medium, large) and location (on the table, inside experimental area excluding those on the table, and outside the experimental area).

User Experience: We used the User Experience Questionnaire (UEQ) [41] to measure the quality of the participants' user experience in each condition.

AR Tabletop Gaming Questions: We designed additional custom questions about different aspects of the VH and the experiment and asked participants to choose their preferred condition and explain their choice (see Table 1).

4.3.3 Behavioral Measures

During the experiment, the participants' head position and orientation tracked by the HoloLens' internal tracking system were logged. From the tracking data, we extracted the following measures.

Head Motion: We measured the amount of overall head motion of the participant by calculating the length of the trajectory the participant's gaze (forward vector) traveled on a unit sphere that surrounds the head (i.e., the origin of the forward vector) during the game, and divide it by the duration of the game.

Dwell Time Ratio on VH: The ratio of time devoted to looking at the VH during the game. We computed an angle between the forward vector and a vector from the user's eye toward the center of VH's head, and we counted the time in which the angle was below 10 degrees.

Dwell Time Ratio on Token: The ratio of time devoted to looking at the VH's token during the game. We computed an angle between the forward vector and a vector from the user's eye toward the center of the token, and we counted the time in which the angle was below 10 degrees.

4.3.4 Hypotheses

Based on the related work and our study design, we formulated the following hypotheses:

H1 Participants indicate higher co-presence with the VH when they observe its ability to move a physical token ($C_P > C_V$).
H2 Participants indicate a more enjoyable gaming experience when the VH can move a physical game token ($C_P > C_V$).
H3 Participants transfer their experience of the VH being able to move a physical token on the table to other physical objects.
H4 Participants exhibit different gaze behavior in the C_P condition compared to the C_V condition.

TABLE 1
AR tabletop gaming related questions.

O1	In which condition did you feel that you were playing a tabletop game with another person?
O2	In which condition did you feel that the virtual human was able to handle physical game pieces?
O3	In which condition did you enjoy the game more?
O4	Would you like to have such a tabletop gaming system at home? Which one would you prefer?

4.4 Results

This section presents the results of the subjective and behavioral measures in the experiment.

4.4.1 Subjective Measures

The questionnaire responses were analyzed using Wilcoxon signed-rank tests at the 5% significance level. Pair-wise comparisons were conducted between the physical and virtual token conditions. We performed multiple comparisons with Bonferroni correction for the object categories in the perceived physicality questionnaire. Box plots in Figure 4 are in Tukey style with whiskers extended to cover the data points which are less than $1.5 \times$ interquartile range (IQR) distance from 1st/3rd quartile.

Co-Presence: The results for the CPQ questionnaire [40] are shown in Figure 4(b). As is common practice for this standard questionnaire, we computed the mean of all ratings from questions 1 to 8 with an inverted score for question 4 (Cronbach's α = .894). We found a significant difference between the two conditions (Z = -2.923, $p = 0.003$), indicating a higher sense of togetherness when the VH can move a physical token.

Perceived Physicality: The results for this measure are shown in Figure 4(c). We computed the means of the ratings for all the objects in each group. In this measure, higher scores indicate that participants rated the VH's ability to move physical objects in this condition higher. As expected, when comparing the physical and virtual token conditions we found significantly higher ratings in the condition with the physical token for the small objects ($Z = -3.060$, $p = 0.002$), medium objects ($Z = -2.488$, $p = 0.013$), large obejcts ($Z = -1.965$, $p = 0.049$), objects on the table ($Z = -2.956$, $p = 0.003$), objects in the experimental area ($Z = -2.440$, $p = 0.015$), and objects outside ($Z = -2.371$, $p = 0.018$). Looking at the physical token condition in more detail, we compared the effect that seeing the VH move a small physical token on the table had on the participants' sense that the VH could move other objects in the room (in the absence of direct evidence for or against this ability). We found a significantly higher probability for participants to judge that the VH could move an object on the table than anywhere else in the room ($Z = -3.408$, $p = 0.003$) or outside the room ($Z = -3.416$, $p = 0.003$). We further found a significantly higher probability for participants to judge that the VH is able to move a small object than a medium ($Z = -3.409$, $p = 0.003$) or large object ($Z = -3.185$, $p = 0.003$).

AR Tabletop Gaming Questions: At the the end of the experiment participants were asked the custom questions in Table 1. Based on their responses, we categorized them in four groups which were *physical*, *virtual*, *both*, and *none*. Figure 4(d) shows the number of participants in each group for each question. We further converted the responses into two columns per question, each column representing the physical and virtual conditions, with *1* assigned for the chosen and *0* for the unchosen. We performed Wilcoxon signed-rank tests on each question. The results showed significant differences between the physical and virtual conditions in participants' responses for O2 ($Z = 4.2$, $p < 0.001$) and O3 (Z = 1.961, $p = 0.05$), indicating a higher physicality expectation for the VH, and a higher enjoyment, in the physical condition. However, we found no significant differences between the conditions, in the answers for O1 ($Z = 0.816$, $p = 0.414$) and O4 ($Z = 1.279$, $p = 0.201$).

User Experience: The results for the UEQ questionnaire [41] are shown in Figure 4(a). For this standard questionnaire, means and variances for all 26 questions are computed between -3 and 3, with scores higher than 0.8 indicating a more positive evaluation. We found a significant difference between the means in the two conditions (Z = -3.002, $p = 0.003$), indicating a higher user experience when the VH could move the physical token.

4.4.2 Behavioral Measures

For the analysis of the behavioral data, we performed paired-samples t-tests at the 5% significance level for each measure. Results for all behavioral measures are shown in Figure 5.

Head Motion: Participants moved their head significantly more in the C_V condition (M = 0.185 m/s, SD = 0.045) than in the C_P condition (M = 0.169 m/s, SD = 0.043); $t(31)$ = -2.341, $p = 0.026$.

Dwell Time Ratio on VH: We found a significant difference in the time participants dwelled on the VH between the C_P condition (M = 0.316, SD = 0.204) and the C_V condition (M = 0.239, SD = 0.153); $t(31)$ = 2.504, $p = 0.018$. Participant spent more time looking at the VH in the physical token condition than with the virtual token while playing the game.

Dwell Time Ratio on Token: We found a significant effect of the conditions on the time participants dwelled on the physical (M = 0.191, SD = 0.144) or virtual (M = 0.260, SD = 0.108) token; $t(31)$ = -2.808, $p = 0.009$. Participants looked down at the VH's token more in the C_V condition than in the C_P condition.

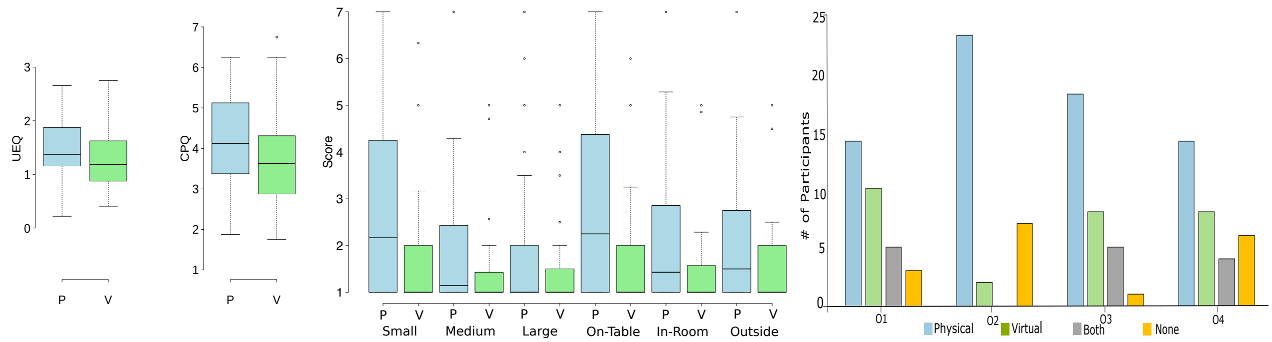
4.5 Discussion

Overall, the sense of co-presence with a VH as well as the perceived physicality of the VH and the user experience was greatly increased by observing the VH's ability to physically affect users' space. In contrast, participants' behavior seemed to be more affected by the limitations of the current state AR glasses, while their gaze behavior showed the potential of our physical-virtual table in mitigating the limitations. In the following, we discuss the results of the experiment in depth, provide potential explanations and implications.

4.5.1 Increased Co-Presence

Our results indicate that the sense of co-presence with the VH was significantly higher in the physical token condition where the VH exhibited its ability to affect the user's physical space compared to the virtual-only condition. The results support our Hypothesis H1.

Our findings are in line with a recent study by Kim et al. [3], in which participants reported a higher level of co-presence with a VH that walked towards a lamp (showing awareness of physical entities) and performed a plausible manipulating gesture to turn on the lamp (showing the ability to affect physical entities) compared to a VH that used a non-physical means to complete the task. The VH in both conditions in our experiment exhibited a similar level



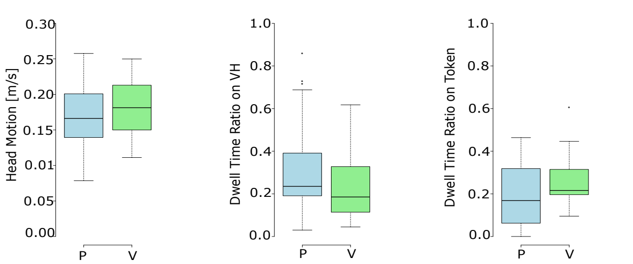
(a) UEQ (b) CPQ (c) Physicality (d) Gaming Questions

Fig. 4. Subjective results with *P* and *V* indicating the physical and virtual token conditions, respectively: (a) user experience questionnaire, (b) co-presence questionnaire, (c) perceived physicality questionnaire with higher scores indicating a higher perception of the VH's ability to move physical objects, and (d) numbers of participants indicating preferences grouped based on their answers to each AR tabletop gaming question.

of awareness of the surrounding physical space, i.e., the VH moved her token to the designated spots on a physical game board, looked at the number wheel on the small monitor at the side of the table, and looked towards the participant when it was their turn. Hence, the increased sense of co-presence in the physical condition is likely mainly a result of the VH's ability to *affect* the physical space and less of the *awareness* of the physical space in our study.

4.5.2 Increased expectation of VH's ability on other objects in the physical space

Regarding the perceived physicality, our results show a significant effect that participants were more likely to believe that the VH would be able to move other physical objects when they observed the VH move the physical token on the tabletop surface, thus supporting our Hypothesis H3. However, it is interesting that the participants were less likely to expect the VH to be able to move objects of larger size than the small physical token or when the distance of the object from the location of their observation of the VH's physical influence increased. When we asked participants about the criteria for their answers, we noticed that most of our participants applied criteria to the virtual human they would also apply to a real human. For example, one participant said "because she could move the real token, she also can move small objects," and another participant explained it with "the size of the object and how heavy it is." In other words, participants expected the VH to behave like a real human and have physical abilities in line with a real human. Although our measure of perceived physicality might limit the interpretation of the physical ability to manipulating real objects, we believe that observing one aspect of physicality—e.g., manipulating a real object—would affect other aspects of physicality, e.g., sensing capabilities of the VH as in [3]. Along these lines, it is also interesting to note that one participant mentioned to have paid more attention to the VH's actions in the virtual token condition because the VH was perceived to be able to cheat more easily with the virtual token than with the physical token.



(a) (b) (c)

Fig. 5. Results of the behavioral measures with P and V indicating the physical and virtual token conditions, respectively: (a) head motion, (b) dwell time ratio on VH, (c) dwell time ratio on token. Whiskers in the box plots are extended to represent the data points less than 1.5 IQR distance from 1st and 3rd quartile.

4.5.3 Improved the user experience of AR game

The UEQ questionnaire is designed to assess user experience in terms of attractiveness, perspicuity, efficiency, dependability, stimulation and novelty [41], which are important elements of an engaging game. The subjective responses for this UEQ questionnaire, the game-related questions listed in Table 1, as well as the informal feedback collected from our participants all are in support of our Hypothesis H2 that the physical token condition would result in a more enjoyable experience. Many of our participants described their interaction as fun, interesting, and exciting. It should be noted that it appears that the limited field of view of the HoloLens may have worked in advantage of the physical token condition, since it satisfied the efficiency and dependability aspects of the UEQ more than the virtual condition according to some of our participants.

4.5.4 Mitigated the usability issue of small augmented FoV

The results for the behavioral measures partially support our Hypothesis H4. We found significant differences between the two conditions in participants' head motion behavior (amount of head motions, dwell time on VH, dwell time on token) in favor of the physical token condition. These differences could be caused by the relatively small augmented field of view of the HoloLens used in this study. Similar to what was described in a recent paper by Lee et al. [29], participants in our study could not see both the VH's face and the virtual token at the same time during the game. Thus, they needed to keep moving their head up and down to see the progress of the game as well as maintain the social interaction with the VH. Whereas, for the physical condition, they could just look down with their eyes to check the position of the opponent's physical token while keeping their head up. Once participants observed the VH's hand touching and moving the physical token, they could mentally connect the VH's visible upper body behavior with the moving physical token seen in the unaugmented periphery of the HoloLens. The reduced dwell time on the token and increased dwell time on the VH in the physical condition seems to match this explanation. Considering the weight of current-state AR glasses, reducing the amount of required head motion to keep track of large virtual content in close proximity of the user could greatly improve the user experience. In this regards, participants' strong preference of the physical condition, as well as the highly rated user experience, might to some degree result from the reduced head motion. AR glasses with a wide augmented FoV can reduce the head motion required to look at the virtual token. However, users' gaze behavior may still be different between two conditions due to the increased co-presence and increased expectation of VH's ability in physical condition. In such a case, an actual eye-tracking should be used instead of the heuristics used here.

4.5.5 Limitations/Potential of the Physical-Virtual Table

The apparatus presented in Section 3 showed a reasonable performance as indicated by the aforementioned high sense of physical-virtual interactivity judged by the participants in our experiment. During the debriefing, when asked about the potential cause of the physical token's movement, 10 participants described it with terms such as *mechanical*, *external force*, or *motorized*, while 15 participants described it as *magnetic*. The fact that most participants came up with a potential computer-controlled cause of the physical movements might be related to the overall high level of computer expertise among our participant population. It would be interesting to compare our results in this experiment to children and participants with less computer experience in future work.

A limitation of the current realization of the prototype is the humming background noise by the motors of the translation stage. During the debriefing, when asked whether they heard sounds while playing the game, 25 participants stated that they did not perceive any noise related to the movement of the token, while 9 participants perceived some noise coming from the table and/or token. In our study, we used headphones to compensate for the background noise of the system, but for future realizations of such actuator systems for tabletop gaming and related experiences, we suggest integrating a noiseless translation stage.

Overall, 23 participants indicated that they enjoyed the condition with the actuated physical token more than the virtual condition, and 18 participants indicated that they would like to have such a tabletop gaming system with actuated physical game tokens at home. We believe that tabletop mechanical actuator systems as described in this paper have much potential for a wide range of tabletop gaming scenarios including serious games such as strategic or tactical wargaming scenarios, e.g., based on an AR Sand Table (ARES) [42] and related efforts.

5 Experiment 2

In this section, we further investigate the characteristics of the observed latency in the actuated surface and the effects of latency on users' perception of observing the virtual human moving a physical token, i.e., physical-virtual interaction.

5.1 Latency in Physical-Virtual Interaction

Two third of the participants in Experiment 1 reported that they observed a lag between the motion of the virtual hand and that of the physical token—eight of them judged the delay as moderate or higher. The observed lag in our actuated surface setup mainly results from the network delay when the optical tracking software transmits the tracked position of the physical token to the AR application running on the HoloLens and the smoothing technique we used to filter out the noise in the tracked position when updating the virtual hand position. As a result, the virtual hand motion was slightly delayed compared to the physical token.

This *physical-to-virtual* latency is specific to our implementation. However, in contrast to updating the virtual hand position to match the translation stage, one could update the position of the translation stage to match the virtual hand position, in which the direction of data transmission would be reversed compared to the implementation we used in Experiment 1. In such a case, the token's motion would be slightly delayed compared to the virtual hand, i.e., causing a *virtual-to-physical* latency. In the following, we consider physical-to-virtual latency as a latency with a negative sign compared to virtual-to-physical latency. For instance, when we talk of a latency of -200 ms we mean a physical-to-virtual latency of 200 ms, i.e., the virtual hand is 200 ms behind the physical token.

In this experiment, we analyze how participants perceive the magnitude and directionality of the latency in the mixed reality tabletop setup.

5.2 Participants

We recruited 13 participants (6 female, 7 male, age 19–56, average 29.8) from the local university community for this study. All of the participants had normal or corrected vision; 5 participants wore glasses and 2 participants wore contact lenses during the experiment. None of the participants reported known visual disorders. Participants used a 7-point scale (1=no expertise to 7=expert) to rate their level of familiarity with VR (average 5.2), AR (average 4.8), VHs (average 4.6), and tabletop games (average 5.2). 10 participants ranked their level of computer expertise as proficient or expert.

5.3 Material

We used the same physical-virtual table setup described in Section 3 but modified the mechanism of synchronizing the hand motion and token motion in order to study how the direction and magnitude of latency affect users' perception of the observed physical-virtual interaction, i.e., the virtual human moving the physical token.

We first simplified the token motion to travel the game board once with stopping only at the corners—four motion

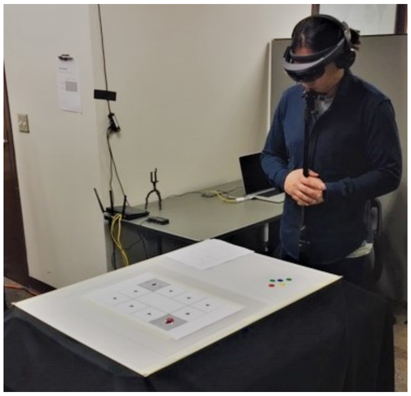


Fig. 6. Experimental setup for Experiment 2. Participants stood at a side of the table and kept their head position fixed during the experiment.

commands were sent at once through a serial communication, then the micro-controller of the translation stage executed the commands in order. For the hand motion, we recorded the entire sequence of the token motion and played it back instead of updating the hand position based on the physical token position in real-time. By doing so, we could initiate the hand motion and token motion separately at different points in time.

To compensate for the network delay, we triggered the hand motion first with a fixed delay (3 seconds) then triggered the token motion with an adjusted delay (3 seconds minus the measured network delay). Then, we varied the order of the initiations of the two motions as well as the waiting time between two initiations by adding or subtracting a target delay time. We prepared 15 latency conditions from -350 ms to 350 ms in steps of 50 ms.

However, it should be noted that there are a few factors that we do not have control over. The hidden internal process of the translation stage generates an arbitrary latency between the trigger and the actual start time of the motion. Also, Unity's Invoke method which we used to trigger the motion with a delay has a small varying offset. We measured the offset between the token motion from a reference motion recorded for the hand motion as well as the offset from the delayed trigger, and applied a post-hoc adjustment to correct the latency conditions shown to the participants.

5.4 Method

We used a within-subjects design. Each participant observed the virtual human moving the physical token around the game board 30 times in total. We prepared two sets of 15 latency conditions, from - 350 ms to + 350 ms in steps of 50 ms, and randomized the order between participants.

5.4.1 Procedure

Upon arrival, participants were given a study brief, protocol, and informed consent. Once they agreed to participate in the study, they donned the HoloLens and the headphones and went through the Interpupillary Distance (IPD) calibration procedure with help from the experimenter and then were guided to stand in front of the table. Participants were asked not to move their head during the experiment and we provided them with a chin rest (see Figure 6). During the experiment, participants were asked to look at the VH's hand and the physical token when they were in motion; otherwise, they were asked to look at the VH. The VH looked at the participants by default, while she moved her gaze toward the token before and while moving the token. Once the VH completed moving the token, participants answered four questions we prepared using verbal responses. At the end of every six observations, participants had a short break and were asked to move one token from the left to the right side. For this, we placed 5 tokens on the participant's side of the table. Upon completion of 30 observations, they took off the HoloLens and headphones and took a survey containing demographics and open-ended questions.

5.4.2 Measures

We prepared four representative questions. Each item intended to measure the overall perceived realism, causality, co-presence, and latency, respectively. Participants rated each question on a 7-point Likert scale. For realism and causality, we asked participants to rate how much they agree or disagree with the following statements (1=strongly disagree, 7=strongly agree): for realism, *"The virtual human's movement of the token seemed realistic"*; for causality, *"The token was moved by the virtual human's hand"*. For co-presence, participants answered the question, *"How much did it seem as if you and the virtual human you saw were together in the same place?"* (1=not at all, 7=very much). And for the perceived latency, participants were asked to choose the delay category they saw (1=the token movement was extremely delayed, 4=the token and hand moved together, 7=the hand movement was extremely delayed).

5.4.3 Data Preparation

Due to the random latency factors in our actuated surface setup (see Section 5.3), we recalculated the occurred latency for each observation using the measured offsets. Then, based on the adjusted latency values, we regrouped the data into six groups:

H-: -300 ms to -200 ms,
M-: -200 ms to -100 ms,
L-: -100 ms to 0 ms,
L+: 0 ms to +100 ms,
M+: +100 ms to +200 ms,
H+: +200 ms to +300 ms.

For the sake of convenience, we refer to the groups with indicators High (H), Moderate (M), and Low (L). A positive sign indicates a virtual-to-physical latency, i.e., the physical token was delayed behind the virtual hand. A negative sign indicates a physical-to-virtual latency, i.e., the token's motion preceded the virtual hand motion.

5.4.4 Hypotheses

The real-world counterpart of the event participants observed in this experiment has a strong causal relationship, i.e., a hand moves a token. Violating the temporal order of the cause and effect might break the illusion of copresence with respect to the simulated event, if any was induced. For example, if the token moved before the virtual human approached it, users would hardly perceive this as a plausible interaction, though there also might be a tolerable delay. Based on this rationale and our study design, we formulated the following hypotheses:

H1 Participants indicate lower causality, co-presence, and realism regardless of the sign of the latency when the magnitude of the latency is high.

H2 Participants indicate higher causality, co-presence, and realism when the physical token's motion is slightly delayed compared to the hand.

5.5 Results

This section presents the results of the subjective measures and the range of delay participants rated as "no delay".

The questionnaire ratings were pooled and analyzed using Friedman tests at the 5% significance level. Wilcoxon signed-rank tests with Bonferroni correction were performed for post-hoc comparisons of the groups. Results for each measure are shown in Figure 7. Significant main effects of latency were found on the perceived causality ($\chi^2 = 52.910$, $p < 0.001$), perceived realism ($\chi^2 = 58.055$, $p < 0.001$), and perceived co-presence ($\chi^2 = 49.584$, $p < 0.001$). Results of the post-hoc comparisons are shown in Table 2.

For the perceived latency ratings, we asked participants to choose the delay category they perceived per each observation during the experiment. We grouped the adjusted latencies by each category (see Figure 8).

5.6 Discussion

Overall, our results show a strong effect of latency on the perceived causality of synchronized physical and virtual motions, overall realism of the observed physical-virtual interaction, as well as the sense of co-presence with a VH affecting physical space. The results are in line with our Hypothesis H1, suggesting negative effects of high latency on the perception of the physical-virtual interaction, independent of the sign of the latency.

Moreover, the results indicate an overall higher tolerance and even a preference for a slight amount of virtual-to-physical (positive) latency compared to the opposite direction, which is in line with our Hypothesis H2. Participants rated higher scores for all measures when the token's motion was slightly delayed compared to the virtual hand.

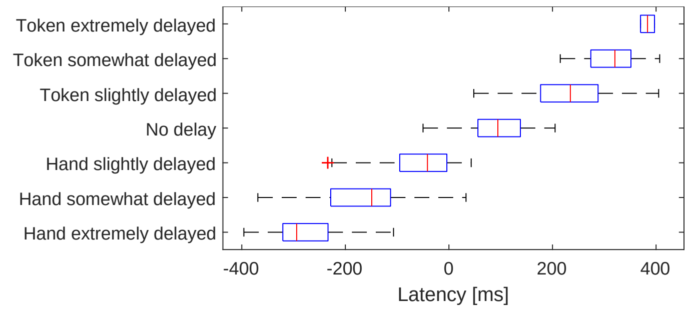


Fig. 8. Ranges of latencies based on the perceived delay categories. Whiskers in the box plots are extended to represent the data points with less than 1.5 IQR distance from 1st and 3rd quartile.

We believe that this direction of latency is more tolerant in terms of preserving a natural causal relationship between a hand and an object that is moved by the hand, considering friction and similar effects that are known to play a role in similar situations in the real world. We received mixed comments on this effect from our participants. One participants commented, *"virtual hand being slightly ahead seems about as good as virtual being slightly behind,"* while one commented, *"Virtual hand moving ahead of the token seemed more realistic than the physical token moving ahead of the hand even if there was a slight delay present."* Meanwhile, our results for the tested range of latencies indicate that there is a perceptual bias in the perceived amount of latency based on which motion was delayed; participants rated latency group L+ more as "no latency" while L- was rated more as "slight delay."

Psychological studies suggest that people perceive the world not by an instantaneous moment-by-moment construction but by integrating information within a short temporal window, while also suggesting that our brain may fabricate the temporal order of stimuli to make sense, e.g., the causal context, during an integration [43]. Given this, it is interesting that our participants' prior knowledge of the causal relation may have shaped their perception of the physical-virtual interaction. Overall, we seem to be more tolerant toward the virtual-to-physical latency direction as we often observe such a delay between a cause and effect in

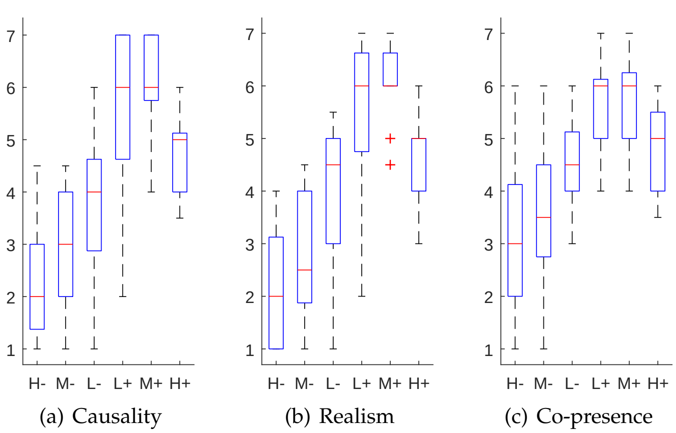


Fig. 7. Subjective results for each group (see Section 5.4.3). Whiskers in the box plots are extended to represent the data points with less than 1.5 IQR distance from 1st and 3rd quartile.

TABLE 2
Summary of the pair-wise Wilcoxon signed-rank test results. Adjusted alpha level = .003 was used to determine significance (*).

Comparison	Causality		Realism		Co-Presence	
	Z	p	Z	p	Z	p
H- vs. M-	-2.295	.022	-1.916	.055	-1.150	.250
H- vs. L-	-3.070	.002*	-3.078	.002*	-2.615	.009
H- vs. L+	-3.192	.001*	-3.186	.001*	-3.068	.002*
H- vs. M+	-3.194	.001*	-3.192	.001*	-3.069	.002*
H- vs. H+	-3.195	.001*	-3.194	.001*	-2.809	.005
M- vs. L-	-2.536	.011	-2.958	.003*	-2.384	.017
M- vs. L+	-3.202	.001*	-3.187	.001*	-3.072	.002*
M- vs. M+	-3.194	.001*	-3.188	.001*	-3.065	.002*
M- vs. H+	-2.982	.003*	-2.989	.003*	-2.623	.009
L- vs. L+	-3.197	.001*	-3.213	.001*	-2.969	.003*
L- vs. M+	-3.077	.002*	-3.192	.001*	-2.956	.003*
L- vs. H+	-2.148	.032	-1.845	.065	-1.406	.160
L+ vs. M+	-1.552	.121	-1.703	.088	-1.512	.131
L+ vs. H+	-1.558	.112	-1.671	.095	-2.714	.007
M+ vs. H+	-2.808	.005	-3.219	.001*	-2.840	.005

the real world. The results of pair-wise comparisons for the moderate and high delay groups (M− vs. M+, H− vs. H+) are in line with this rationale, thus supporting H2.

However, although the overall patterns are similar, it seems that the sense of co-presence is less affected by the latency compared to the causality or realism. For instance, one participant commented: *"[..] the perspective rendering, proper occlusion, clipping (via field of view), etc. played a large role in establishing the answer to this question regardless of the delay."*

6 Conclusion

In this paper, we investigated the effects of a virtual human's physical influence on participants' perception of the virtual human and its abilities. We described an apparatus based on a motorized translation stage capable of magnetically moving small physical objects over a tabletop surface, while the physical source of the movement is hidden from an observer's view. Instead, in this setup, users wear a HoloLens and see a virtual human reach out with its hand and move the physical object. Based on this setup, we designed a basic interaction scenario, a tabletop board game, and performed a user study where participants played the game twice, each time with the virtual human either moving a virtual or a physical token throughout the game. Our results show significant benefits of the virtual human being able to move a physical token with respect to a positive impact on participants' sense of co-presence, physicality, and the virtual human's abilities.

We further addressed the research question of how the latency between physical and virtual movements in this mixed reality setup affects the perceived plausibility of the interaction with the virtual human. We formalized the latency in two directions with the physical object's movement preceding the virtual hand's movement or vice versa. Our results show that a slight temporal delay of the physical token moving after the virtual hand lead to a significant increase in ratings of realism, co-presence, and perceived causality during the mixed reality interaction.

Future work may focus on extending the presented setup to the third dimension, i.e., moving physical objects not only on the tabletop surface but integrating an electromagnetic mechanism to levitate them in mid air (e.g., see [39]). This would enable situations where the virtual human could pick up an object from the tabletop and set it down again, such as when picking up and rolling dice.

Acknowledgments

The work presented in this publication is supported in part by the Office of Naval Research (ONR) Code 34 under Dr. Peter Squire, Program Officer (ONR awards N00014-14-1-0248, N00014-12-1-1003, and N00014-17-1-2927). The work is also supported in part by the National Science Foundation (NSF) under Grant Number 1800961 (Dr. Tonya Smith-Jackson) and 1564065 (Dr. Ephraim P. Glinert). We also acknowledge AdventHealth for their support of Prof. Welch via their Endowed Chair in Simulation.

References

[1] K. Kim, M. Billinghurst, G. Bruder, H. Been-Lirn Duh, and G. F. Welch, "Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017)," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 24, no. 11, pp. 2947–2962, 2018.

[2] S. Orts-Escolano, C. Rhemann, S. R. Fanello, D. Kim, A. Kowdle, W. Chang, Y. Degtyarev, P. L. Davidson, S. Khamis, M. Dou, V. Tankovich, C. Loop, Q. Cai, P. A. Chou, S. Mennicken, J. Valentin, P. Kohli, V. Pradeep, S. Wang, Y. Lutchyn, C. Keskin, and S. Izadi, "Holoportation: Virtual 3D Teleportation in Real-time," in *ACM Symposium on User Interface Software and Technology (UIST)*, 2016, pp. 741–754.

[3] K. Kim, L. Boelling, S. Haesler, J. N. Bailenson, and G. B. G. F. Welch, "Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the Perception of Intelligent Virtual Agents in AR," in *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2018, pp. 105–114.

[4] M. Porcheron, J. E. Fischer, S. Reeves, and S. Sharples, "Voice Interfaces in Everyday Life," in *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2018, pp. 640:1–640:12.

[5] J. Blascovich, "Social influence within immersive virtual environments," in *The Social Life of Avatars*, ser. Computer Supported Cooperative Work, R. Schroeder, Ed. Springer London, 2002, pp. 127–145.

[6] J. Blascovich, J. Loomis, A. C. Beall, K. R. Swinth, C. L. Hoyt, and J. N. Bailenson, "Immersive virtual environment technology as a methodological tool for social psychology," *Psychological Inquiry*, vol. 13, no. 2, pp. 103–124, 2002.

[7] C. Harms and F. Biocca, "Internal Consistency and Reliability of the Networked Minds Measure of Social Presence," in *Annual International Presence Workshop*, 2004, pp. 246–251.

[8] J. Fox, S. J. Ahn, J. H. Janssen, L. Yeykelis, K. Y. Segovia, and J. N. Bailenson, "Avatars versus agents: A meta-analysis quantifying the effect of agency on social influence," *Human-Computer Interaction*, vol. 30, no. 5, pp. 401–432, 2015.

[9] N. Magnenat-Thalmann, G. Papagiannakis, and P. Chaudhuri, "Interactive virtual humans in mobile augmented reality," in *Encyclopedia of Multimedia*. Springer, 2008, pp. 362–368.

[10] M. E. Hoque, M. Courgeon, J. Martin, B. Mutlu, and R. W. Picard, "Mach: My automated conversation coach," in *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2013, pp. 697–706.

[11] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis, "Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments," *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 6, pp. 583–598, 2001.

[12] J. Llobera, B. Spanlang, G. Ruffini, and M. Slater, "Proxemics with multiple dynamic characters in an immersive virtual environment," *ACM Transactions on Applied Perception (TAP)*, vol. 8, no. 1, 2010.

[13] F. Argelaguet Sanz, A.-H. Olivier, G. Bruder, J. Pettre, and A. Lecuyer, "Virtual Proxemics: Locomotion in the Presence of Obstacles in Large Immersive Projection Environments," in *Proceedings of IEEE Virtual Reality (VR)*, 2015, pp. 75–80.

[14] S. T. Bulu, "Place presence, social presence, co-presence, and satisfaction in virtual worlds," *Computers & Education*, vol. 58, no. 1, pp. 154–161, 2012.

[15] E. Goffman, *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York, NY USA: The Free Press (a Division of Simon and Schuster, Inc.), 1963.

[16] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst, "Mini-me: an adaptive avatar for mixed reality remote collaboration," in *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2018, pp. 46:1–46:13.

[17] M. Dragone, T. Holz, and G. M. O'Hare, "Using mixed reality agents as social interfaces for robots," in *Proceedings of the IEEE International Symposium on Robot and Human interactive Communication (RO-MAN)*, 2007, pp. 1161–1166.

[18] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[19] S. C. Khullar and N. I. Badler, "Where to look? automating attending behaviors of virtual human characters," *Autonomous Agents and Multi-Agent Systems*, vol. 4, no. 1-2, pp. 9–23, 2001.

[20] B. Scassellati, "Mechanisms of shared attention for a humanoid robot," in *Embodied Cognition and Action: Papers from the AAAI Fall Symposium*, vol. 4, no. 9, 1996.

[21] R. A. Brooks, C. Breazeal, R. Irie, C. C. Kemp, M. Marjanovic, B. Scassellati, and M. M. Williamson, "Alternative essences of intelligence," *AAAI/IAAI*, vol. 1998, pp. 961–968, 1998.

[22] C. Breazeal, "Toward sociable robots," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 167–175, 2003.

[23] A. P. Saygin, T. Chaminade, H. Ishiguro, J. Driver, and C. Frith, "The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions," *Social Cognitive and Affective Neuroscience*, vol. 7, no. 4, pp. 413–422, 2012.

[24] J. Hochreiter, S. Daher, G. Bruder, and G. Welch, "Cognitive and touch performance effects of mismatched 3d physical and visual perceptions," in *Proceedings of IEEE Virtual Reality (VR)*, 2018, pp. 379–386.

[25] A. Kotranza and B. Lok, "Virtual human+ tangible interface= mixed reality human an initial exploration with a virtual breast exam patient," in *Proceedings of IEEE Virtual Reality (VR)*, 2008, pp. 99–106.

[26] P. Lincoln, G. Welch, A. Nashel, A. Ilie, A. State, and H. Fuchs, "Animatronic shader lamps avatars," in *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2009, pp. 27–33.

[27] M. Obaid, I. Damian, F. Kistler, B. Endrass, J. Wagner, and E. André, "Cultural Behaviors of Virtual Agents in an Augmented Reality Environment," in *International Conference on Intelligent Virtual Agents*, ser. Lecture Notes in Computer Science, Y. Nakano, M. Neff, A. Paiva, and M. Walker, Eds., vol. 7502. Springer Berlin Heidelberg, 2012, pp. 412–418.

[28] E. Kruijff, J. E. Swan, and S. Feiner, "Perceptual issues in augmented reality revisited," in *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2010, pp. 3–12.

[29] M. Lee, G. Bruder, T. Höllerer, and G. Welch, "Effects of unaugmented periphery and vibrotactile feedback on proxemics with virtual humans in ar," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 4, pp. 1525–1534, 2018.

[30] K. Kim, G. Bruder, and G. Welch, "Exploring the effects of observed physicality conflicts on real-virtual human interaction in augmented reality," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 2017, pp. 1–7.

[31] M. Sra, S. Garrido-Jurado, C. Schmandt, and P. Maes, "Procedurally generated virtual reality from 3d reconstructed physical space," in *Proceedings of the ACM Conference on Virtual Reality Software and Technology (VRST)*, 2016, pp. 191–200.

[32] A. L. Simeone, E. Velloso, and H. Gellersen, "Substitutional reality: Using the physical environment to design virtual reality experiences," in *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2015, pp. 3307–3316.

[33] Z. He, F. Zhu, and K. Perlin, "Physhare: Sharing physical interaction in virtual reality," in *Adjunct Publication of the ACM Symposium on User Interface Software and Technology (UIST)*, 2017, pp. 17–19.

[34] S. Kasahara, R. Niiyama, V. Heun, and H. Ishii, "extouch: Spatially-aware embodied manipulation of actuated objects mediated by augmented reality," in *Proceedings of the ACM International Conference on Tangible, Embedded and Embodied Interaction (TEI)*, 2013, pp. 223–228.

[35] J. Lifton, M. Laibowitz, D. Harry, N. Gong, M. Mittal, and J. A. Paradiso, "Metaphor and manifestation cross-reality with ubiquitous sensor/actuator networks," *IEEE Pervasive Computing*, vol. 8, no. 3, 2009.

[36] M. Lee, K. Kim, S. Daher, A. Raij, R. Schubert, J. N. Bailenson, and G. Welch, "The wobbly table: Increased social presence via subtle incidental movement of a real-virtual table," in *Proceedings of IEEE Virtual Reality (VR)*, 2016, pp. 11–17.

[37] S. Follmer, D. Leithinger, A. Olwal, A. Hogge, and H. Ishii, "inform: dynamic physical affordances and constraints through shape and object actuation," in *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2013, pp. 417–426.

[38] D. Leithinger, S. Follmer, A. Olwal, and H. Ishii, "Physical telepresence: shape capture and display for embodied, computer-mediated remote collaboration," in *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2014, pp. 461–470.

[39] J. Lee, R. Post, and H. Ishii, "Zeron: mid-air tangible interaction enabled by computer controlled magnetic levitation," in *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2011, pp. 327–336.

[40] C. Basdogan, C. Ho, M. A. Srinivasan, and M. Slater, "An experimental study on the role of touch in shared virtual environments," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 7, no. 4, pp. 443–460, 2000.

[41] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Applying the user experience questionnaire (UEQ) in different evaluation scenarios," in *Proceedings of the International Conference of Design, User Experience, and Usability*, 2014, pp. 383–392.

[42] C. R. Amburn, N. L. Vey, M. W. Boyce, and J. R. Mize, "The Augmented REality Sandtable (ARES)," US Army Research Laboratory, Tech. Rep. ARL-SR-0340, 2015.

[43] H. Choi and B. J. Scholl, "Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception," *Perception*, vol. 35, no. 3, pp. 385–399, 2006.

Myungho Lee is a postdoctoral researcher at the Korea Institute for Science and Technology. He received his M.S. and Ph.D. in Computer Science from the University of Central Florida, in 2016 and 2019 respectively, under the supervision of Prof. Gregory F. Welch. He received his M.S. from Korea University in 2009, and B.S. in Electrical Engineering from Ajou University in 2007. His research interests include virtual/augmented reality, human perception, human-computer interaction.

Nahal Norouzi received her bachelor's degree from Electrical Engineering in Amirkabir University of Technology in 2015. She is a Computer Science PhD student in the University of Central Florida and a Graduate Research Assistant in the Synthetic Reality (SREAL) Lab. Her research interests are Augmented and Virtual reality, Human-Computer Interaction, and Agent Perception.

Gerd Bruder is a Research Assistant Professor at the Institute for Simulation and Training at the University of Central Florida. He received his Habilitation in Computer Science from the University of Hamburg in 2017, his Ph.D. in Computer Science from the University of Münster in 2011, and his M.Sc. in Computer Science with a minor in Mathematics from the University of Münster in 2009.

Pamela J. Wisniewski is an Assistant Professor in the College of Engineering and Computer Science at the University of Central Florida. Her research interests are situated in Human-Computer Interaction and lie at the intersection of social computing and privacy. Her goal is to frame privacy as a means to not only protect end users, but more importantly, to enrich online social interactions that individuals share with others.

Gregory F. Welch is a Professor and the AdventHealth Endowed Chair in Simulation at the University of Central Florida, with appointments in the College of Nursing, the Department of Computer Science, and the Institute for Simulation & Training, and is a Co-Director of the Synthetic Reality Laboratory. He is also an Adjunct Professor of Computer Science at the University of North Carolina at Chapel Hill. His research interests include human-computer interaction, human motion tracking, virtual and augmented reality, computer graphics and vision, and training related applications.