Blended Agents: Manipulation of Physical Objects within Mixed Reality Environments and Beyond

Susanne Schmidt  
Universität Hamburg  
Germany  
schmidt@informatik.uni-hamburg.de

Oscar Ariza  
Universität Hamburg  
Germany  
ariza@informatik.uni-hamburg.de

Frank Steinicke  
Universität Hamburg  
Germany  
steinicke@informatik.uni-hamburg.de

ABSTRACT

Mixed reality (MR) environments allow real users and virtual agents to coexist within the same virtually augmented physical space. While tracking of different body parts such as the user’s head and hands allows virtual objects to show plausible reactions to actions of the real user, virtual agents only have a very limited influence on their physical environment.

In this paper, we introduce the concept of blended agents, which are capable of manipulations of physical properties related to the object’s location and surface material. We present two prototypic implementations of virtual-physical interactions using robotic actuators and thermochromic ink. As both interactions show considerably different characteristics, e.g., with regard to their persistence, explicability, and observability, we performed a user study to investigate their effects on subjective measures such as the agent’s perceived social and spatial presence. In the context of a golf scenario, participants were interacting with a blended agent that was capable of virtual-physical manipulations such as hitting a golf ball and writing on physical paper. A statistical analysis of quantitative data did not yield any significant differences between blended agents and VAs without physical capabilities. However, qualitative feedback of the participants indicates that persistent manipulations improve both the perceived realism of the agent and the overall user experience.

CCS CONCEPTS

• Human-centered computing → Mixed / augmented reality; Empirical studies in HCI.

KEYWORDS

Virtual Agents; Physical Interaction; Mixed Reality Environments

1 INTRODUCTION

In modern mixed reality (MR) environments, the links between real users and virtual agents (VAs) are bidirectional in many regards: By means of advanced display technology, VAs can be rendered as three-dimensional spatial entities within the same environment as the user, while head tracking allows agents to also detect and react to the user’s position within this environment. Natural language processing enables VAs to understand their communication partners while speech synthesis and natural dialogue systems generate human-like responses. By contrast, modern sensing technology such as head or hand tracking systems allows virtual objects to show physically correct reactions to actions of the real user, while the manipulation of real objects through VAs is only possible in a very limited scope, e.g., in the form of coherent global illumination. Other virtual-physical interactions such as simulated collisions between VAs and physical objects need more complex actuators, in particular if the technology behind this interaction should be hidden from the user to create an advanced illusion of plausible human-like agents. This additional complexity is leading to an asymmetry between real and virtual interaction partners as illustrated in Figure 1.

As a consequence, only few examples of virtual-physical interactions can be found in the literature (examples are discussed in Section 2.5). Instead, most of the current MR applications accept implausible physical-virtual collisions to a certain extent or try to avoid them when possible [13, 15].

In this paper, we introduce the concept of blended agents which are able to manipulate real-world objects in an interactive way. Throughout the paper, we focus on two different forms of virtual-physical interactions:

1. Manipulations of physical properties related to the object’s location.
2. Manipulations of physical properties related to the object’s surface material.

To address the first form of virtual-physical interactions, we utilized an off-the-shelf robotic golf ball that moves along a scripted path to simulate interaction with a virtual golf player. For the second form, we designed a novel device that uses temperature variation to activate thermochromic ink on a sheet of paper. In the presented prototypic setup, a pre-defined score appears virtually before it is replaced by a physically persistent version. Synchronized with the animations of a blended agent, the illusion of a virtual human writing on a physical piece of paper can be created.

Both forms of virtual-physical interactions not only differ from each other with regard to the underlying physical change, but may also cause varying user responses due to differences in what we call...
explicability, observability, and persistence. In terms of explicability, we assumed that manipulations of an object’s position achieved via motors or magnetic actuators are both more common and, depending on the implementation, more conspicuous as they might also influence other properties of the prepared object such as its weight. Changing an object’s surface material, however, usually requires chemical reactions, e.g., to temperature variations, pressure, or UV light. Such chemical changes are uncommon in other application fields and usually do not interfere with other object properties. If users are not capable of finding an obvious explanation for an effect, this might support the illusion of an interaction between the blended agent and the physical object. Furthermore, during the MR experience, manipulations of an object’s position can be observed directly, while changes of the surface material might not be easily detectable by users as similar effects can be achieved by overlaying virtual projections. Therefore, we hypothesize that during the MR experience changes of the material might be not recognized by the user at all. On the other hand, since such changes can be persistent they could be observed by the user outside of the MR environment. Whether such a long-term manipulation of real-world objects could change the perception of the VA’s realism retrospectively is one of the questions we intended to answer in a user study.

Taking these considerations into account, we formulate the following hypotheses:

(H1) Virtual-physical interactions improve the user experience in terms of social and spatial presence, ecological validity, perceived anthropomorphism of the blended agent, and engagement.

(H2) Virtual-physical interactions related to the surface material of an object have a stronger positive impact on the aforementioned metrics than those related to the object’s position.

(H3) Chemical changes of an object’s surface material can be hidden from the user, while mechanical manipulations of the object’s location are more explicable.

(H4) Manipulations of the object’s position are observed by the users directly, while manipulations of the object’s surface material are not observed before the end of the MR experience.

To our knowledge, no prior work has investigated similar manipulations and their effects on agent-human interaction.

In summary, the contributions of this paper are:

- Implementation of an API for an off-the-shelf robotic ball to simulate collisions with a virtual golf club.\(^1\)
- Development of a proof of concept for a thermal table that allows blended agents to persistently write on physical sheets of paper.
- Collection of subjective quantitative and qualitative user responses to compare both forms of virtual-physical interactions in a between-subjects design.

The remainder of this paper is structured as follows. Section 2 gives an in-depth view of the state of the art considering different dimensions of VAs’ realism. In the context of virtual-physical interactivity of virtual agents, we designed and implemented two different forms of manipulations that are presented in Section 3. In Section 4 we present a human-subject study that compares the effects of both interactions on the perceived anthropomorphism of blended agents as well as the elicited sense of spatial and social presence. Section 5 concludes the paper and discusses future research.

\(^1\)https://github.com/augmentedrealist/spheromini.js
2 BACKGROUND

The realism of a VA was initially considered as a synonym for its human-like visual appearance. However, advanced models from research fields such as speech synthesis, motion capture, and non-verbal communication led to an extended definition that involves VAs with natural language, realistic behavior and attention towards their environment as well as their human communication partners. This conceptual change raises interesting research questions regarding the correlation between the manifold dimensions of human likeness of VAs and social factors such as perceived copresence in a shared MR environment. In the following, we summarize a selection of previous studies that investigated this correlation, grouped by the dimension of human likeness they considered.

2.1 Agent Embodiment

Though personal digital assistants have become widespread in the context of smart homes as well as professional environments, most of the current implementations rely on audio output or displayed text only. By means of VR/AR technology, such voice-based VAs can be supplemented with a humanoid virtual body. A number of research projects addressed the question whether and how agent embodiment affects the social interaction between VAs and real humans.

In the context of museums, Schmidt et al. [24] showed benefits of virtual embodied guides over those implemented just by voice. Visitors of a simulated exhibition reported a higher sense of spatial and social presence in the company of embodied VAs. Furthermore, agent embodiment showed a significant effect on the perceived credibility of the VA, indicating that embodied guides appeared to be more competent and trustworthy.

Similar positive effects of agent embodiment on the users’ sense of trust, social richness, and social presence with the VA could be found in a human-subject study by Kim et al. [12]. In addition, participants of the study reported an increased confidence in the agent’s ability to influence the real world and to react to real-world events, when the VA was embodied and showed natural social behaviors.

A literature review on early embodied VAs was presented by Dehn and van Mulken [7]. Their meta-analysis revealed some inconsistent findings regarding the effects of embodied VAs on user experience. While some of the analyzed empirical studies report benefits of embodied VAs, others conclude that agent embodiment only showed little or even negative effects on the users’ responses. Dehn and van Mulken hypothesize that these different outcomes may be attributed to varying degrees of the agent’s appearance, both in terms of visual fidelity and natural voice.

2.2 Agent Appearance

By means of continuous advancements in technology, real-time AR applications can make use of increasingly realistic VAs, both with regard to the visual appearance and quality of synthesized speech. While a positive correlation of a VAs fidelity and the elicited sense of anthropomorphism seems to be reasonable at first glance, several studies demonstrate that the agent’s appearance cannot be considered in isolation, since it strongly correlates with other indicators of agent realism.

For instance, Bailenson et al. [2] investigated the effects of visual and behavioral realism of VAs on perceived copresence. They conclude that both types of realism should be considered in conjunction as large disparities between them resulted in lower levels of copresence. These results are consistent with a previous study conducted by Garau et al. [9], which also revealed a significant interaction effect between agent appearance and behavior. A mismatch between visual fidelity and behavioral realism might also explain an observation made by Nowak and Biocca [22]. To their surprise, VAs with a higher level of anthropomorphism caused a decrease in reported copresence and social presence. The authors argue that anthropomorphic VAs may raise expectations about their behavioral realism and should only be used if the system is able to meet these expectations.

2.3 Behavioral Realism

As already discussed in the previous section, realistic behavior, including natural gestures, body, and eye movements, as well as lip syncing, is a crucial factor for VAs to be perceived and treated as if they were human.

A study of Gratch et al. [10] revealed that even simple non-verbal reactions to the user such as gaze shifts or head nods can cause feelings of rapport. Moreover, Demeure et al. [8] showed that appropriate emotional verbal and non-verbal behaviors of VAs can evoke a higher sense of perceived believability, competence, and warmth. Further studies with a focus on objective measures found that users maintain greater distance from VAs who engage them in mutual eye contact [1], and that culturally inconsistent gaze behavior of VAs results in higher heart rates [23]. As a result of a user study on collision avoidance behavior in small-scale virtual environments, Bönsch et al. [5] recommend introducing an ‘awareness zone’ around VAs. In this zone, users expect VAs to react to their presence by gazing towards them and stepping aside when more space is needed to pass by.

2.4 Environmental Awareness

A special type of behavioral realism relates to the degree to which VAs are aware of their physical environment (for a review see [21] or [11]). Different approaches are possible to endow a VA with knowledge about the physical world, e.g., extracting information from a pre-populated database, or analyzing dynamic sensor data.

The embodied agent MACK [6], for example, is using a static knowledge base including the VA’s fixed location and orientation, as well as the layout of the physical building in which it is located. Based on this information, MACK is able to provide context-sensitive and spatially referenced information, such as directions to a specific room in the MIT Media Lab. Barakonyi et al. [3] developed a framework for autonomous AR agents, which are able to monitor and thus react to changes of real-world attributes. The presented AR Puppet cannot only avoid collisions with physical obstacles, but is also able to support a user in a physical construction task by tracking the user’s progress. In a recent study, Kim et al. [14] demonstrated that by responding to subtle environmental events such as the airflow of a real fan, VAs can create a higher sense of copresence. Despite the positive results of their study, the authors also state that the VA’s awareness behavior was less effective than
techniques that involve active participation of the user, such as the wobbly table experience [17]. In the latter project, Lee et al. introduced the concept of virtual-physical interactivity, which will be considered in detail in the following section.

2.5 Virtual-Physical Interactivity
In AR environments, the creation of a human-like VA turns out to be even more challenging than in VR environments. Even with a maximum degree of visual fidelity, natural voice, realistic behavior, and environmental awareness, the created illusion can suddenly break if the agent is not following the same laws as its physical surrounding. Potential physicality conflicts include unnatural occlusions between the VA and physical objects as well as implausible physical-virtual collisions.

Negative implications of such conflicts on human-agent interactions were demonstrated by Kim et al. [13, 15]. In their studies, the participants observed a VA encountering a physical obstacle such as a door or a chair. In different conditions the VA either avoided collisions with the obstacle, asked the participant to move it out of the way, or passed through the obstacle. Subjective responses indicate that physical-virtual conflicts reduced the sense of copresence while proactive behavior asking help from the users to avoid implausible conflicts increased the ratings of copresence.

Instead of avoiding collisions between VAs and physical objects, another approach is to allow VAs to actually interact with these objects, e.g., to move them to a different location. In this context, we introduce the concept of blended agents - VAs that are not only capable of influencing their virtual surrounding but also of performing virtual-physical interactions.

The latter concept was first investigated by Lee et al. [18]. They implemented a tabletop game that can be played by a real human and a VA. Via an actuator system underneath the surface of the table, the VA is able to move not only virtual tokens but also a physical token. In a within-subjects study, the authors were able to show benefits of the VA moving a physical token both with regard to subjective and some behavioral measures. In this condition, participants reported a higher sense of copresence, physicality, and the agent’s abilities.

Another example of VAs that are capable of influencing their physical environment was implemented by Lee et al. [17]. Their custom-made Wobbly Table crosses the boundary between the physical and virtual world. While the physical half is standing in front of a projection screen, a virtual counterpart is visually extended into the virtual environment with the VA. If the VA is leaning on the table, a virtual-physical interactions occurs as both the virtual and the physical part are slightly tilted.

In a recent paper, Lee et al. [16] demonstrated that even subtle tactile footstep vibrations induced via the floor can increase subjective estimates of social presence in an AR environment.

While all of these research projects already demonstrate the potential of blended agents for enhancing MR experiences, there are still a number of open research questions. Firstly, each of the implemented virtual-physical interactions is based on manipulations of the physical object’s pose. As the range of physical changes covers many other effects, e.g., related to the object’s shape and surface material, it would be interesting whether the previously observed effects are generalizable to these forms of physical manipulations and if there are individual differences between them.

In order to compare manipulations related to the object’s pose and surface material we implemented two setups that are detailed in the next section. Secondly, all of the presented studies rely on a within-subjects design, therefore allowing the participants a direct comparison of agents with and without virtual-physical capabilities. As users in real-world applications usually do not have this comparison and virtual-physical manipulations are not yet common enough to be assumed to be the norm, the question arises whether users expect VAs to be capable of physical manipulations to appear human-like. Therefore, in the second part of the paper we present a between-subjects study that aims to measure unbiased reactions to virtual-physical interactions.

3 Apparatus
Our primary goal in this project was to gain insight into MR experiences that involve different forms of interactions between blended agents and physical objects. For this purpose, two different setups were implemented to exemplify virtual-physical interactions in the form of (i) movements of a physical ball, and (ii) writing on a physical sheet of paper.

3.1 Robotic Ball
To create the illusion that a blended agent is moving a physical ball, we utilized an off-the-shelf Sphero mini. This motorized ball has a diameter of 42 mm and therefore matches the size of a customary minigolf ball. The locomotion system as well as additional sensors are hidden inside an outer shell. Two motor-driven wheels move along the inner surface of the shell to actuate a weight at the opposite side (see Fig. 2). Due to the resulting weight shift the ball can move with three degrees of freedom (i.e., rotation around the y axis and x/z translation) with a maximum speed of 1 mps. During navigation, the ball is stabilized using an inertial measurement unit (IMU) that contains both an accelerometer and a gyroscope.

While we could take advantage of the mechanics of the Sphero mini, the control from within a game engine required some implementation effort. Since at the time of writing this paper no SDK for this Sphero model was available, we implemented a custom JavaScript library based on existing SDKs of the preceding models. The library contains commands to establish a connection to the robotic ball and to control its motors. The code is executed on a dedicated NodeJS server, which can receive motion commands from any Unity3D application via HTTP requests. The commands are translated to Sphero Mini machine code and sent to the device via Bluetooth LE. The full source code has been uploaded to GitHub.

3.2 Thermal Table
To address virtual-physical manipulations on a material level, we did a lot of research and testing to find a dye that is invisible to the eye under normal conditions and can be activated by external stimuli such as light or temperature. Photochromic dyes turned out to be unsuitable for MR environments with real users as they usually have to be exposed to high-intensity UV light in order to change their color within seconds. We therefore decided in favor of the thermochromic ink and evaluated four different activation temperatures...
(-10°C, 31°C, 47°C, and 90°C) in advance of the user study to find the best tradeoff between practicability and user experience. The inks with the highest and lowest activation temperature featured an irreversible color change but had to be discarded, nonetheless, as touching the activating thermal element creates a strong sensation of cold/heat, which may reveal the hidden technology and even pose a danger to the uninformed user. For this reason, we narrowed down the choice to the two inks with medium activation temperatures and reversible behavior. To still convey the user the impression of a permanent writing, we finally chose the 47°C dye,

SFXC thermochromic color changing screen ink for paper and board, Black 47C.

For the activation of the thermochromic ink, we built a thermal device composed of a Peltier element that can display temperature variation (see Fig. 3). The device features the TEC1-12706 thermoelectric plate, it measures 40x40mm, operates from 0~15V and 0~6A, and the temperature range goes from -30°C to 70°C. We used the plate in conjunction with a heat sink and a fan, as a mechanism to dissipate the heat from the bottom side. In order to variate the temperature on the top side, to be a cold or a hot end, we use an H-Bridge circuit that switches the polarity of the voltage applied. As a result, we can transfer heat with a cooling rate of 4°C/s and a heating rate of 8°C/s.

The plate is installed on the top of a pedestal (1.0m), offering access to the temperature-switching side of the Peltier device with a tilt angle of 22.2°. An Espressif ESP32 controls the thermoelectric plate, a dual-core microprocessor clocked at 240 Mhz with 520Kb RAM and powered by rechargeable 18650HG2 Li-ion batteries (3.7V). All the electronic components are mounted inside the pedestal and connected to the control PC using a USB cable. The cable handles the serial communication with the device at (57Kbps), enabling the PC to send three commands in order to enable/disable the thermoelectric plate, switch the temperature from high (65°C) to low (10°C) or vice-versa, and activate the thermo-active ink accordingly.

In the current implementation of the thermal table as a proof of concept, all written text has to be prepared before the MR experience. Only when placed on the powered table, the prepared text turns invisible until the polarity of the thermoelectric plate is changed and the sheet of paper is cooled down. In Section 4.6, we will discuss some thoughts on possible enhancements of the setup to allow flexible writing instead of pre-defined text only.

4 USER STUDY

As the proposed systems implement two different forms of virtual-physical interactions, one of whom similar to previously tested techniques and one that features a novel approach, we conducted a comparative study to collect subjective responses of users. For this purpose, we designed an experimental environment that naturally embeds both implementations, as described in detail in the following section.

4.1 Materials

The MR environment for the user study was inspired by a minigolf course, with a VA acting as the opposing player. The course with a length of 2.9 meters and a width of 1.05 meter was set up within a four-sided CAVE, which featured a size of 4.2 × 3.13 × 2.36m. Two heavy ropes marked the edges of the course and pieces of artificial turf served as obstacles. The hole was marked with a slightly raised ring. Three Optoma EH320USTi short throw projectors were used to display virtual content at the CAVE walls. Two additional Optoma GT1080(e) projectors augmented the floor as well as the thermal table. Each projector provided a resolution of 1920×1080 at a refresh rate of 120 Hz. To experience the view-dependent stereoscopic content, users had to wear shutter glasses with passive markers that were tracked by a five-camera OptiTrack system. Furthermore, the voice of the VAs was presented to participants via wireless noise-cancelling headphones. Another purpose of the headphones was to block ambient noise that was created by the thermal table’s internal fan. In contrast, sounds caused by the friction between golf ball and floor were still audible.

All of the actions of the participants were monitored by the experimenter using a camera at the ceiling of the CAVE. By this
means, the experimenter was able to trigger particular reactions of the blended agent from a neighboring room, without being visible to the participants. The CAVE setup was chosen since current AR headsets only feature a small field of view and virtual objects are cut off at the edges of the display area.

As we learned from related research projects, consistency between different dimensions of realism seems to be crucial for the perceived human likeness of VAs. For this reason, the VA used in the experiment has to meet several expectations regarding his appearance, speech, and behavior. For a high degree of visual fidelity, we used a 3D scanned female head model that features a highly detailed mesh, 4K PBR textures, and a number of natural facial expressions.\(^3\) To add vividness to the VA’s face, her eyes were not moving randomly but were focusing occasionally on special points of interest such as the golf ball or the user. Furthermore, micro movements, i.e., saccades, as well as blinking reflexes were performed. The agent’s body was created and rigged using Adobe Fuse as well as Mixamo. As retargeted keyframe animations were described as stiff and artificial in a pre-study, we decided to replace them with motion captured material. All animation sequences were performed by a female actor and recorded with an 8-camera Qualisys Miqus M3 system. A path of the robotic golf ball was programmed accordingly to match the animations of the blended agent. In addition, a native speaker provided the voice of the VA, and matching lip movements were created via the Oculus Lip Sync plug-in. To improve the realism of sound propagation, we used the Unity MS HRTF spatializer plug-in, which incorporates the binaural head-related transfer function. Besides the voice, no additional sound effects were included in the current experiment.

### 4.2 Methods

For the user study, we followed a between-subjects design with two independent variables and two levels each. The resulting four conditions, all in relation to the VA’s interactions, are:

- \((B_v H_v)\) Virtual golf ball and virtual handwriting.
- \((B_v H_r)\) Virtual golf ball and real handwriting.
- \((B_r H_v)\) Real golf ball and virtual handwriting.
- \((B_r H_r)\) Real golf ball and real handwriting.

The random assignment of conditions was counterbalanced among participants.

Before a new participant arrived, the golf course was prepared according to the selected condition. In the conditions \((B_v H_v)\) and \((B_v H_r)\) the thermal table was turned on and a sheet of paper prepared with invisible ink was placed on its top. In preparation to the conditions \((B_r H_v)\) and \((B_r H_r)\), the NodeJS server was started and a connection of Unity3D to the robotic ball was established. Since the Sphero is not able to provide a global orientation value, the initial rotation of the golf ball had to be determined by hand. A manual correction was performed until the ball moved perfectly along a test track. Afterwards the ball was positioned at its starting slot along with three other physical golf balls.

Before they entered the previously described minigolf course, participants had to fill in a consent form as well as a pre-questionnaire to provide demographic information. Afterwards, each participant was guided to the CAVE and the procedure as well as general minigolf rules were explained. In this introductory phase, participants were able to examine the golf course as well as the scorecard with their naked eye. Also, the preparation of the scorecard with a table was executed in sight of the participants to make sure that they realize it was empty when they entered the room. After all questions were resolved, the participant had to wear shutter glasses and the experimenter started with the first round.
In total, four rounds of play were performed:

1. The experimenter playing with a physical ball.
2. The participant playing with a physical ball.
3. The (blended) agent playing with a virtual/robotic ball.
4. The participant playing with a physical ball.

After rounds (1) and (2), the experimenter and participant filled in one blank of the scorecard each. Afterwards, participants were introduced to the VA by the experimenter. They were given noise cancelling headphones and the experimenter left the room. The VA then started a conversation with the participant and putted either a virtual or a robotic ball into the hole, according to the selected condition (see Fig 4b). After finishing round (3), the VA walked to the scorecard and asked the participant where to fill in her score, as illustrated in Figure 4a. The animation was only resumed by the experimenter, if the participant was close to the table. This artificial pause should ensure that all participants witnessed the handwriting of the VA and therefore do not make false assumptions about how and when the VA's score was added to the scorecard. If the participant was in sight of the thermal table, the VA virtually wrote a pre-defined score of 4 in the dedicated blank space. The VA then challenged the participant to bet his score in round (4). During this second round of the participant, all hits were counted by the experimenter using the live camera view. In the conditions (B_vH_r) and (B_rH_v), when the participant was close to the hole, the temperature of the thermal table was switched from high to low and the thermochromic ink below the projected score became visible. At the same moment, the projected score was faded out with the result that the participant could only see the physically written score when he returned to the table.

After round (4) was finished and the participant filled in the last blank of the table, the VA started a final evaluation of the match. Based on the number of strokes that were digitally logged by the experimenter, the VA announced the winner of the game. Finally, the VA said goodbye and suggested the participant to take the scorecard as a souvenir. The user left the CAVE and was asked to fill in a series of questionnaires. Overall, the study took around 20 to 25 minutes to complete.

4.3 Participants
We invited 40 participants to our study, 27 male and 13 female (aged from 18 to 41, M=25.35). 36 of them were students or staff members of the local department of computer science, while 4 stated to pursue a non-technical profession. According to the pre-questionnaire, 7 participants took part in a study involving VR or AR for the first time. None of the 40 participants reported any visual impairments that could affect the results of our experiment.

4.4 Results
During the user study, we collected both quantitative and qualitative subjective data that can give some indication of how different virtual-physical interactions affect a MR experience. The results are presented in the following section.

4.4.1 Quantitative Analysis. Each experiment session was concluded with five questionnaires that addressed different aspects of the experience:

- Social presence
  (= Social Presence Questionnaire by Bailenson et al. [1])
- Spatial presence
  (= subscale of the Temple Presence Inventory [20])
- Ecological validity
  (= subscale of the ITC Sense of Presence Inventory [19])
- Perceived anthropomorphism
  (= subscale of the Godspeed questionnaire [4])
- Engagement
  (= top loading items of the engagement subscale of the ITC Sense of Presence Inventory [19])

Results were measured on 7/5-point Likert scales, as noted in the head of Table 1. For each participant, average scores were formed according to the computation models that are suggested in the original papers. We analyzed the data using multiple two-way ANOVAs, but could not find any significant main or interaction effects. The mean values as well as standard deviations for all conditions and each dependent variable are also summarized in Table 1.

4.4.2 Qualitative Analysis. In addition to the mentioned Likert scales, we were also asking participants some open-ended questions about experienced or expected effects of blended agents, depending on the tested condition:

(B_v) “How did the agent’s interaction with a real golf ball affect your experience?”
(B_r) “Imagine the agent interacting with a real golf ball instead of a virtual one. How would this interaction have affected your experience?”
(H_v) “How did the agent’s persistent handwriting affect your experience?”
(H_r) “How would your experience have been changed, when the handwritten score of the agent would be still visible on the paper?”

We directed similar questions at participants assigned to the virtual and the real conditions as we were interested in the users' perception of individual potentialities of both virtual and blended agents. By this means, we also aimed to investigate whether expectations for virtual-physical interactions and the reality diverge to some extent. To extract comparable data, we assigned utterances to seven different categories using an open coding strategy. The first three categories denote opinions regarding the perceived realism of blended agents in comparison to virtual agents without virtual-physical capabilities. Another four categories cover different dimensions of emotional responses. If a single response of a participant included multiple utterances within the same category (e.g., “fascinating and memorable”), they were still counted only once. As each scenario (B_v, B_r, H_v, H_r) was experienced by 20 participants, 20 is the maximum value in each category. The resulting frequency distribution is illustrated in Table 2.

In addition to the categorized utterances, four participants acknowledged the increased fairness when both the user and the blended agent have to play with a physical golf ball. Regarding the scorecard, four of the participants with a (H_r) condition reported their initial surprise when the projected score disappeared. One participant even admitted to feel disappointed as he could not take the completed scorecard as a souvenir. In general, three participants...
mentioned that a scorecard with physically written text feels more like a trophy.

We were also interested whether reactions to the persistent handwriting were different for users assigned to a (B_r) or (B_v) condition. Experiencing a blended agent that is interacting with a real golf ball might have raised the expectations on the agent’s capabilities. However, no such interaction effect could be found as both user groups showed similar, mainly positive, responses to the persistent handwriting.

Finally, we asked participants of the (B_r) or (H_r) conditions about the used mechanism in order to test our hypothesis (H3). For the physical golf ball, 8 of the participants stated that they figured out the mechanism or at least got an idea of how it worked. Surprisingly, only one of the ideas was correct while most participants suspected a magnetic track behind the ball movement. In contrast, none of the (H_r) participants perceived the mechanism behind persistent handwriting as obvious. Due to a lack of explanations, two participants were convinced that another person entered the room to replace the virtual score, while another two felt uncertain about the fact whether the scorecard was empty at the beginning of the study.

4.4.3 Observational Data. In addition to feedback obtained through the questionnaires, we also made some observations regarding the participants’ behavior, both during the study and directly afterwards.

When the agent was approaching the scorecard to fill in the blank, six of the participants (five of H_r and one of H_v) used their own pen to write down the agent’s score in her stead.

After the MR experience but before completing the post- questionnaires, all participants of the conditions (B_r,H_r), (B_v,H_v), and (B_r,H_r) were asked whether they noticed the physicality of the golf ball and handwriting, respectively. 16 participants who experienced a B_r condition realized that the ball was real during the experiment, while 4 reported to have doubts whether the ball was real or not. In contrast, only 7 participants of the H_r conditions noticed that the handwritten score was still persistent after they left the MR environment. These results support our initial hypothesis (H4) regarding the observability of both forms of virtual-physical interactions. It was surprising, though, as we expected that users will realize the persistent handwriting at the latest when they leave the MR environment.

4.5 Discussion

Against the in (H1) and (H2) hypothesized positive effects of blended agents on several social and spatial factors, no significant differences could be found in the collected data. These statistical results can be interpreted in different ways. Two possible implications might be that (i) there actually are no differences between VAs without virtual-physical capabilities and blended agents, and (ii) some users react positively to blended agents while others show negative reactions, compensating one another. In contradiction to both approaches to an explanation are the qualitative comments that were collected at the end of the study. The majority of participants...
reported a positive influence of both virtual-physical interactions in terms of perceived realism and/or user experience. The question arises why the quantitative ratings do not reflect these subjective impressions. In the following, we discuss several potential influencing factors and rate their respective impact.

**Limited Expectations on VAs** Starting point of the following discussion is the basic question "Do users expect VAs to have physical capabilities?". The fact that the majority of participants did not notice the persistent handwriting at all and showed emotional responses of surprise and confusion when they were made aware of it is indicative of rather low expectations on the VA. Therefore, users assigned to a virtual condition most likely compared the displayed VA to agents they experienced in the past, without a negative impact of missing physical capabilities. This impression is supported by the qualitative feedback as users of the virtual condition felt positive about the realistic body movements, the natural voice, and individual reactions of the VA. Therefore, the reported effects found in previous within-subjects studies might be only due to the direct comparison between agents.

**Low Granularity of Used Scales** Three of the questionnaires used 5-point Likert scales as we complied with the standards. As current VAs are still far from being indistinguishable from real humans, only few users will rate items that are related to the human likeness with a maximum value of 5. Therefore, only one of the remaining options refers to a positive response. A higher scale granularity that allows participants to rate their experience more precisely might reveal significant differences between the conditions. That these differences are expected to be rather small was already indicated by results of a similar study with a within-subjects design [18]. For the ratings of engagement an additional ceiling effect can be observed as mean scores are already close to the maximum for conditions without blended agents.

**Limited Importance of the Physical Reactions** Although both the golf ball and the scorecard were designed to be an integral part of the interaction between the participants and the MR environment, they might have been less meaningful than other interactions with physical objects. For example, if a blended agent moves a real chair towards the user to take a seat, this physical manipulation has an impact on the subsequent actions, while the physical golf ball could only be observed without any direct contact. In another example, a blended agent could mark a location on a physical map to direct the user to a place. In contrast, the scorecard was only given as a souvenir without any future purpose. More meaningful interactions might have increased the perceived value of the physical (persistent) manipulations.

**Distrust of the Experimental Environment** An observation that might have influenced the results without being the sole reason was that some participants conjectured that somebody entered the room and replaced the virtual score by a physical one when the participant was distracted by the golf match. Even the fact that the experimenter was neither in the CAVE nor in the directly neighboring room could convince them of the contrary. Two other participants mentioned that they were sure that the sheet of paper was empty at the first, but were skeptical about this in retrospect as they would not know how this could have been done.

### 4.6 Limitations

Both the interactions with the robotic ball and the thermal table are proof of concepts with some limitations.

As the robotic ball has no global tracking capabilities its position and orientation has to be determined manually. A computer vision algorithm could be used to compute the current position of the ball and to match subsequent actions of the blended agent. Such a tracking algorithm could also compensate for the limited precision of the Sphero mini. In the current implementation, the robotic ball ends up in slightly varying places. While the blended agent always put the ball into the hole, the golf club and the ball movements were not always perfectly in sync; an observation that was also shared by some of the participants. Another limitation is related to the ball physics. As any motorized objects need an acceleration phase to be set in motion, the initial impulse imparted to the golf ball by hitting it with a golf club cannot be simulated completely.

The current implementation of the thermal table also requires some preparations in order to create a convincing illusion of a blended agent. First of all, the ink can only be made visible at once. This is why the handwritten text has to be prepared before the MR experience. For the same reason, the writing path has to be simulated virtually before it is replaced by the physical writing. To solve both problems, coated paper could be used. In this case the thermal mechanism has to be changed from a heating plate to a heated metal tip as used for soldering irons. Furthermore, as the used thermochromic ink is visible at room temperature, it always has to be placed on top of the thermal table at the beginning of a MR experience. By using a dye with different characteristics users could bring a sheet of paper to the MR environment, which might further increase the believability of the blended agent.

### 5 Conclusion

All actions performed by VAs are usually restricted to computer-generated objects, resulting in an asymmetry between real and virtual interaction partners. In this paper, we investigate the concept of blended agents; VAs that are able to cross the boundary between a virtual and the physical world. We implemented two exemplary manipulations that are affecting real-world objects either within a MR environment (i.e., moving a physical golf ball), or even outside of a MR environment (i.e., writing on a physical sheet of paper).

To compare both forms of virtual-physical interactions we conducted a user study with 40 participants. Although a statistical analysis of subjective data obtained through a number of questionnaires did not yield any significant differences between blended agents and VAs without physical capabilities, user responses still provide insight into the potential of virtual-physical manipulations. Users described their interaction with blended agents as an "amazing, very surprising and immersive experience", a "fascinating magic trick", or the sensation of "being inside the Holodeck". The agent’s physical manipulations made the agent appear more present", and created a "more enjoyable" and "more memorable" MR experience.
In spite of little divergences between natural and simulated behavior of the physical golf ball, the majority of participants appreciated the virtual-physical interaction. Some participants also mentioned effects on their behavior inside the MR environment, as they “felt the urge to respond” the blended agent or avoided any collisions.

To take up these points, we want to collect objective data such as the users’ avoidance behaviors or points of interest in a future study. In terms of subjective responses, we plan to use scales with higher granularity to obtain more differentiated feedback. Furthermore, a future study design should focus on virtual-physical interactions that are more meaningful to the users, e.g., a persistent writing future study design should focus on virtual-physical interactions.

In terms of subjective responses, we plan to use scales with higher granularity to obtain more differentiated feedback. Furthermore, a future study design should focus on virtual-physical interactions that are more meaningful to the users, e.g., a persistent writing future study design should focus on virtual-physical interactions.

REFERENCES


