2022 31st IEEE International Conference on Robot and
Human Interactive Communication (RO-MAN)
August 29 - September 2, 2022. Naples, Italy.

Benefits of an Interactive Robot Character in Immersive Puzzle Games

Ting-Han Lin[*,1], Spencer Ng[*,1] and Sarah Sebo[1]

***Abstract*— Robots are becoming increasingly prominent in the entertainment sphere, where they interact with guests in themed environments to tell stories, often in place of human characters. To evaluate the potential benefits of robots in these contexts compared to humans, we created an interactive puzzle game where either a robot or a human actor serves as a diegetic "game guide" character that is both a cooperative partner and an omniscient game master. In the game, participants solve a crime mystery by asking the game guide for information to complete tasks and for hints to solve puzzles. We conducted a between-subjects study ($n = 42$) to investigate how players' game experiences differed when the game guide was a human compared to an embodied robot. Our results show that participants playing with a robot had more fun, felt less judged, and felt more connected with the robot while solving tasks compared to those playing with a human. These results suggest that robots can be effective alternatives to human actors in broader immersive entertainment contexts such as escape rooms to provide greater enjoyment and promote more social interaction with in-game characters.**

I. Introduction

The entertainment industry is rapidly moving towards providing immersive experiences for guests, where domains such as themed entertainment, theater, and escape rooms tell individualized and interactive stories. These pieces are often mediated through technology for both its novelty effect and its ability to create scalable experiences that have a high throughput [1]. Using technology to tell stories is not a new idea, with animatronic robots found in theme park attractions such as the Enchanted Tiki Room and the Carousel of Progress since 1963 [2]. However, immersion-creating techniques today often involve live interaction, such as in escape rooms where human actors play in-game characters or a robotic Mr. Potato Head telling jokes to guests in Toy Story Mania's queue [3]. Industry trends are thus engaging narrative characters in a participatory culture with guests to tell stories together, rather than characters serving as a mere communication medium to act out scenes.

We hypothesize that robots can be effective in providing engaging immersive experiences in entertainment contexts because people will feel less social judgment from a robot and act more naturally around it compared to a human in the same role. Researchers in past studies have observed that English as a second language (ESL) students are more willing to make mistakes in front of a robot compared to native speakers because they felt less anxious with a robot [4]. Similar to students learning new skills, guests in entertainment spaces may be put into new settings, as they

[*]Both authors contributed equally to this work.
[1]University of Chicago; {tinghan,spencerng,sarahsebo}@uchicago.edu

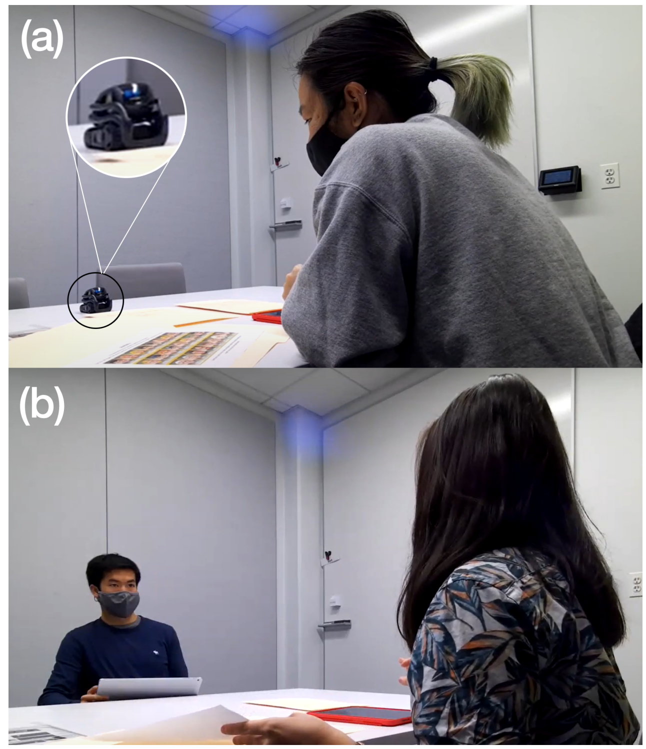

Fig. 1: Participants played with either (a) a robot game guide or (b) a human game guide in a series of games and puzzles.

are expected to role-play and engage in unfamiliar social contexts to experience a story. For instance, *Sorcerers of the Magic Kingdom* is a mixed-reality game where players cast spells while traversing the titular theme park [5], which may lead to feelings of awkwardness due to being seen by visitors outside of the game's "magic circle" who view gameplay as unexpected guest behavior [6]. Introducing robot character companions here may help normalize players' game actions, therefore making them feel more comfortable while facilitating more natural interaction compared to a human character.

Our study investigates the social effects and viability of robots in an immersive live game compared to when players interact with a human in the same capacity. We diegetically place a robot into the game's story, and players must interact with the robot to progress. Although prior studies in human-robot interaction have independently examined how robots can be effective game masters by providing personalized help to players and how robots differ from humans in social presence via short scripted interactions [7], [8], [9], none have integrated robots in a narrative game in which a human is interchangeable with the robot's role. We thus look at how the presence of a robot playing an immersive game with participants may affect their enjoyment and comfort, their task completion, and their perception of the in-game character in contrast with the presence of a human actor.

II. Related Work

A. Interactive Robots in Games and Entertainment

Games have been used in past human-robot interaction work to investigate robots' social attributes, such as team-building [10], facilitating conversations [11], and robot trust [12]. Salomons et al. used a card game to show that human participants tended to conform to robots' answers when creating word-picture associations, particularly when there was a level of trust between participants and the robots [13]. Similarly, Correia et al. demonstrated that humans forming teams with robots while playing a card game tended to have greater trust and identification with the group when robots expressed "group-based" emotions such as using more inclusive pronouns (e.g. "we") to express pride [10]. Card-playing robots also gained players' trust after more play sessions [14]. These studies show the viability of robots in place of humans as sociable and trustworthy game companions, and we adopt techniques from past work when scripting our robot character to connect with players. Nonetheless, the games in these studies are not primarily designed to provide entertainment to players, and we wish to investigate if interactive robots can facilitate enjoyment in games that is comparable to or greater than playing with humans, particularly if tasks are novel and fun even without the robot.

Furthermore, a survey conducted by Muñoz et al. suggests that human-robot interaction researchers should "pay more attention to the story and aesthetics" of games with robots, rather than focusing on game mechanics and technological advances [15]. This finding is important because stories are critical component of games [16], and emergent narrative features requiring player interaction in game environments can lead to increased feelings of immersion and thus enjoyment [17]. Prior work by Kory and Breazeal had used an expressive DragonBot for a storytelling game, where a child and a robot successfully collaborated to tell stories [18]. Ligthart et al. further suggest that children are more engaged and enjoy stories more when they are given the option to develop a story alongside a robot by choosing paths in a branching narrative [19]. We aim to expand upon this framework of agency in our study by providing the illusion of open-ended interaction with a robot while also directly incorporating the robot into the narrative of the game. Therefore, we develop a new type of immersive game where players' enjoyment when playing with a robot has not yet been explored in the literature.

B. Robot Interaction and Social Presence

Past work has effectively introduced robots into various domain roles typically filled by humans. Hayashi et al. showed that two robots performing stand-up comedy have greater social presence compared to recorded human comedians [20]. Similarly, a chess-playing robot was seen as equivalent to a human player in its expressiveness, even if its behaviors were prescripted [21]. These findings suggest that people can view robots more positively than humans in the same social role, even if the robot is primarily passive in responding to users. We thus wish to see if robots can continue to have a larger social presence in narrative games, especially when directly compared to a live human actor, as much of the prior work focuses on pre-recorded or hypothetical human behavior.

Robots as interactive game characters compared to human actors may also induce lesser feelings of social judgment from players who interact with them, which in turn causes players to enjoy the game more. In the context of robots tutoring ESL students, Leyzberg et al. observed that students were more willing to make mistakes when practicing English with the robot when they may not have been as comfortable making those mistakes in front of classmates or their teacher [4]. They note that "young students who are English learners often feel anxiety when interacting with other native English speakers due to various internal and external pressures" and "may perceive the robot differently" compared to a person [4]. Moreover, Bryant et al. observed that children were more engaged when interacting with a robot compared to a human therapist in a virtual reality rehabilitation game, potentially due to higher perceptions of trust and decreased pressure from robots [22]. While this prior work suggests that robots may provide benefits by reducing perceptions of social judgment, no work to our knowledge has investigated whether this difference exists between humans and robots occupying the same gamified tutor role.

III. Methods

In a between-subjects study, participants were placed in a role-playing puzzle adventure with either a robot (robot condition) or human (human condition) game guide. Participants role-played as a detective trying to solve a kidnapping mystery with an in-game character called Agent Lee ("game guide"). Through a series of three puzzle tasks ("dossiers"), participants asked the game guide for the necessary information and optional hints to solve the crime. After completing the tasks, participants filled out a questionnaire about their experience. This study was approved by the University of Chicago's Institutional Review Board (IRB21-1642).

A. Hypotheses

Our study investigated five hypotheses. In Pereira et al., participants playing Risk with a social robot described the experience as "interesting" and "fun," akin to playing against a human [23]. Players said the experience was novel and wanted to play longer with the robot, which we hypothesize will similarly affect players' perception of our study's puzzles when they interact with a robot game guide:

- **H_1**: Playing with a robot character in a game is more enjoyable compared to playing with a human character.

In our game, the game guide acts similarly to a tutor by providing hints. Because Leyzberg et al. suggested students may feel less anxious and be more willing to make mistakes in front of a robot compared to a human tutor [4], students with robot tutors may perform better if they feel more comfortable asking for more help. We hypothesize that the approachable presence and positive perceptions of robot helpers will extend to interactive puzzle games:

- **H**$_2$: Playing games with an embodied robot leads to greater player comfort and reduced social judgment compared to playing with a human.
- **H**$_3$: Participants are more likely to ask for help and thus make greater progress on puzzles when playing with a robot game guide compared to a human.

Finally, we predict that interacting with a "game master" robot will lead players to feel that the robot is more actively engaged than a human in the same role, with respect to giving hints and being attentive to player actions. Leite et al. previously studied children playing chess with an iCat robot and found that participants thought the robot was more expressive when giving advice and more "calm" than a human who may rush the game, despite the iCat being more passive and objective in how it responds to player actions [21]. This shift in perception of the robot's empathy, in conjunction with how meaningful human-robot relationships develop at a similar pace to human-human relationships [24], suggests that robots may be more likable than humans who act in a similar capacity during short-term interactions. Therefore, we hypothesize the following:

- **H**$_4$: Robots acting as in-game characters facilitate stronger interpersonal connection, have a greater social presence, and have lower social expectations from players compared to human actors in the same role.
- **H**$_5$: Robots portraying in-game characters will be perceived as warmer and more sincere in their interactions with players compared to human actors.

B. Conditions

The hypotheses were investigated through two conditions, in which the game guide was played by either a human (human condition) or a robot (robot condition). In both conditions, the game guide gave encouragement and narrative material to participants while responding to participants' questions. The game guide also read from the same script, and the human actor and the robot operator were the same person in the study to maintain consistency in responses.

1) Human Game Guide: The game guide was played by a 20-year-old male-presenting actor who was a trained expert in the study. The human actor sat opposite to the participant and carried a tablet during the study to track their progress through Zoom, keep time, and deliver pre-scripted hints using the interface in Figure 2.

2) Robot Game Guide: The game guide was played by an Anki Vector robot. Vector is a consumer-grade robot that has been marketed as a family pet and used to investigate collaboration and physical play dynamics with children [25], [26], [27]. We chose an embodied robot due to physical presence being a contributing factor towards learning new skills and leading to more positive and natural interactions with humans compared to virtual agents [8], [28].

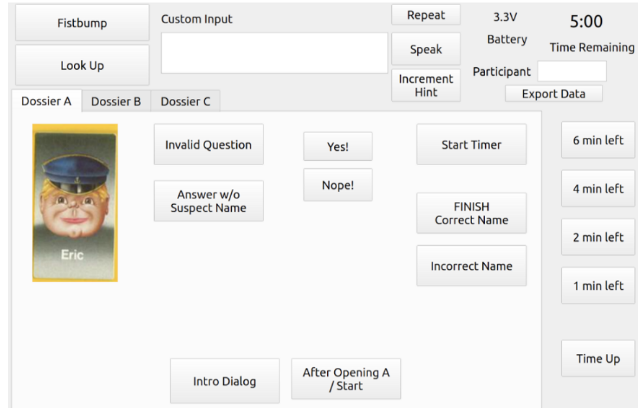

Fig. 2: User interface for Wizard of Oz robot control and delivering hints in both conditions.

The robot was controlled through the Wizard of Oz paradigm using a custom Python interface[1] (Figure 2) that allowed the robot operator to perform the same functions as the human actor, including giving hints, answering questions, and initiating a fistbump. Players were monitored using Zoom through a camera in the study room and the screen sharing feature on the iPad they wrote on. Connectivity to the robot was provided through the Anki Vector Python SDK and the Robot Operating System (ROS). This enabled text-to-speech functionality and motor control to respond to users. We used the Wizard of Oz approach to facilitate an accurate understanding of participant behavior and ensure a similar play experience to the human game guide.

C. Game Design

The game contained three timed puzzles "dossiers" and one tutorial dossier. To solve each dossier, participants needed to use logical reasoning, deduction, and pattern recognition skills. All dossiers had appropriate cluing, such that players could reasonably solve them without asking the game guide for hints (aside from yes/no questions in Dossier A). The dossiers created an immersive environment due to their integration in the story of solving the kidnapping case with the game guide, Agent Lee, and they followed good escape room puzzle design principles as given in [3].

1) Dossier A: The first part of the study had participants playing a "Guess Who?" game, where players must ask Agent Lee yes/no questions (e.g. "Was the suspect wearing a hat?") to narrow down a grid of 24 suspects to one criminal. The task was designed to build players' confidence due to its low difficulty and to familiarize them with the game guide. In Dossier A, the game guide's role was an in-game companion, acting complementary to the player in contrast to being a "puzzle tutor" or "game master" [7]. The game guide thus served as a diegetic interface to get information, which increases immersion in games as suggested by [29].

2) Dossiers B & C: The remaining two tasks were cryptic puzzles, where steps towards progression may not be immediately clear, and multiple steps (e.g. extracting

[1] `https://github.com/SeboLab/interactive-games`

We asked residents where they were when the missing person was last seen and if they saw the suspect. These transcripts don't seem to add up, but can you bring order to them to reveal the location of the suspect's secret hideout?

Alexander: I was at the grocery store, staring at the five electric cars powered by the two solar panels.

Ben: I was at the park on multiple occasions last morning — three I think — and only saw four geese walking there.

Cameron: When I was with my pets (five in total, minus the four mice who don't like the sun), I saw something suspicious at the post office.

Fig. 3: A sample of Dossier B, where participants determined a hideout location with the help of the game guide.

numbers and math operations hidden in sentences) are required to arrive at a solution. The game guide transitioned into a "game master" role that instructed participants and established a sense of intelligence [7]. Participants were also given a Sample Cryptic Dossier and a code sheet (containing conversions between binary numbers, decimal numbers, and letters) to familiarize themselves with the puzzle format. They then solved Dossiers B and C (see Figure 3), which were presented as in-game transcripts to help the player find the suspect's hideout and weapon. Each dossier had four main phases of deduction involved. To create a smooth difficulty curve with respect to the time limit of 10 minutes, Dossier B was more difficult than Dossier C, and the overall experience was designed such that it would be difficult to solve both tasks in time without the game guide's help.

D. Study Protocol

The experimenter led participants into the study room, with thematic music playing in the background. Participants gave informed consent for study participation and were briefed on the story, with the game guide present as either a human or a robot. Next, the experimenter gave a tutorial on asking yes/no questions to the game guide. Participants then played Dossier A for up to 5 minutes and filled out a survey about their experience with Dossier A.

The experimenter then gave a tutorial to participants on how to ask for hints and confirmation on puzzle progress from the game guide using a Sample Cryptic Dossier. We encouraged participants to explicitly ask for help (e.g. "Hey Agent Lee, I need some help.") or ask yes/no questions after stating their progress (e.g. "Hey Agent Lee, I think there are hidden numbers in the words. Am I on the right track?"). After the tutorial, participants played Dossiers B and C with the game guide for 10 minutes, where Dossier B must be solved before moving onto Dossier C. Participants then filled out a survey about their experience with Dossiers B and C, followed by their overall game experience. Upon finishing, the experimenter interviewed participants about their play experience, concluding the study.

During both parts of the study, the game guide only interacted with participants when participants said "Hey Agent Lee" followed by a question. The game guide would also introduce each dossier and its objective while providing time warnings with remaining time. At the end of each part, the game guide gave verbal encouragement and prompted participants to give them a fistbump. Participants were compensated with a $6 Amazon gift card for completing the 30-minute study.

E. Player Interactions with the Game Guide

Pre-written hints were given to participants in Dossiers B and C when they either explicitly asked the game guide for help or for progress confirmation. Each phase of the puzzle tasks had an associated series of hints that led participants to move onto the next phase. The series of hints for a particular phase would become increasingly transparent about what to do, and participants would receive multiple hints for the same phase by asking for additional clarification from the game guide or making incorrect progress checks after receiving an initial hint. For instance, the first hint to solve Phase 2 in Dossier B (finding math operations in transcripts and applying them to the corresponding numbers) suggested that "the two numbers in each transcript could be combined," while the second hint suggested that "there's a hidden math operation in each transcript."

Participants' progress in each phase was tracked through iPad notes and verbalized thoughts, such that the personalized hints given would best help participants solve the task [30]. If participants made an error in their work, the game guide also suggested that they should review their results.

F. Measures

A combination of subjective and objective measures were collected during the study to address our hypotheses:

1) Questionnaire: A Qualtrics survey with various subjective measures was given to the participants in the study. They were asked to rate statements such as "I had fun while solving Dossier A" during each part of the study. They were also asked to rate how much pressure and judgment they felt from the game guide during the overall experience, as well as how "actively engaged" the game guide was.

Each of the statements was on a 7-point Likert scale, with 1 representing that participants "strongly disagree[d]" and 7 meaning they "strongly agree[d]." Participants were also asked to rate the game guide's warmth using the RoSAS subscale [31] and sincerity using the MDMT [32] on 7-point Likert scales, with 1 meaning that certain descriptors were "definitely not associated" with the game guide and 7 meaning they were "definitely associated." Participants' personality traits, background experience solving puzzles, and demographic data were recorded in the survey as potential covariates. At the end of the survey, participants were asked two open-response questions where they described their experiences of interacting with the game guide and how they thought changing the game guide from a human to a robot (or vice versa) would affect their experience.

2) Interview: A short interview asked how participants felt when solving each dossier and what the game guide could do to make the game experience better.

3) Objective Data: Information was collected by the game guide's actor/operator on player performance, including the time to complete each dossier, the number of phases completed in Dossiers B and C, the number of hints asked for and given, and the number of confirmations on progress.

G. Participants

43 participants were recruited from the University of Chicago community via direct recruitment, flyers, and social media. Data from one participant was discarded due to robot malfunction. 31 participants identified as Asian, 10 as White, 2 as American Indian, 2 as Black, and 5 identified as another ethnicity. Participants identified as two or more ethnicities were double-counted. We balanced the gender of participants between our two experimental conditions, beyond which we randomly assigned participants to a condition. 21 participants (11 male, 9 female, and 1 non-binary) played with the human game guide, and 21 participants (11 male, 9 female, and 1 declined to identify their gender) played with the robot game guide. Participants ranged in age from 18 to 23 ($M = 20.00$, $SD = 1.31$), with an average age of 19.90 ($SD = 1.37$) for those in the human condition and 20.52 ($SD = 1.40$) for those in the robot condition.

Using items from the TIPI [33], participants reported personality traits including extraversion ($M = 4.07$, $SD = 0.58$) and openness to new experiences ($M = 4.26$, $SD = 0.61$); those in the human condition rated extraversion as ($M = 5.02$, $SD = 1.17$) and openness to new experiences as ($M = 4.71$, $SD = 1.35$), while those in the robot condition rated extraversion as ($M = 4.45$, $SD = 1.04$) and openness to experiences as ($M = 3.86$, $SD = 0.98$). Participants also reported their familiarity playing cryptic puzzles on a 7-point Likert scale ($M = 3.29$, $SD = 1.77$), where participants showed similar familiarity in the human condition ($M = 3.38$, $SD = 2.01$) and the robot condition ($M = 3.19$, $SD = 1.53$). There were no significant differences among these potential covariates between the conditions.

IV. Results

Quantitative data collected from the study were analyzed for differences using Welch's t-test across the two conditions, and we report effect sizes as Cohen's d. We also present qualitative observations to support our hypotheses.

A. Gameplay Experience

Participants found the overall gameplay experience more fun with a robot game guide ($M = 6.62$, $SD = 0.50$) compared to a human ($M = 6.10$, $SD = 0.83$, $t = 2.48$, $d = 0.76$, $p = 0.019$; see Figure 4). In particular, there is strong evidence showing that playing Dossier A is more fun with a robot ($M = 6.67$, $SD = 0.58$) than a human ($M = 5.76$, $SD = 1.00$, $t = 3.60$, $d = 1.11$, $p = 0.001$). While participants rated their puzzle-solving experience in Dossiers B and C as more fun with the robot ($M = 6.43, SD = 0.75$) than with a human ($M = 6.05, SD = 0.92$), this difference was not significant ($t = 1.47, d = 0.45, p = 0.15$). These results show support for H_1, that the overall experience was more enjoyable with a robot game guide than a human.

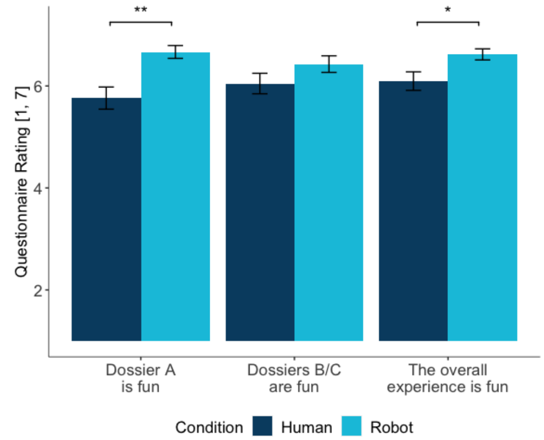

Fig. 4: Participants found solving the dossier tasks to be more fun when playing with a robot game guide compared to a human. (*) denotes $p < 0.05$, and (**) denotes $p < 0.01$. Error bars depict one standard error from the mean.

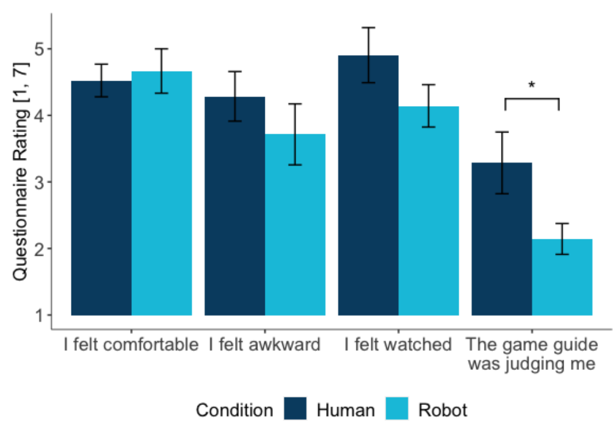

Fig. 5: Participants felt more judged with the human game guide compared to the robot game guide. (*) denotes $p < 0.05$. Error bars depict one standard error from the mean.

B. Player Comfort

Figure 5 depicts how much participants felt certain personal descriptors relating to their comfort during the game experience. Participants perceived that the human game guide was judging them more ($M = 3.29, SD = 2.12$) compared to the robot ($M = 2.14$, $SD = 1.06$, $t = -2.20$, $d = -0.68$, $p = 0.035$), showing support for H_2.

We also found a general trend that participants feel less comfortable, more awkward, and more watched with the human game guide compared to the robot. However, these differences were not significant ($M_{\text{robot}} = 4.67$, $SD_{\text{robot}} = 1.53$, $M_{\text{human}} = 4.52$, $SD_{\text{human}} = 1.12$, $t = 0.35$, $d = 0.11$, $p = 0.73$ for comfortable; $M_{\text{robot}} = 3.71$, $SD_{\text{robot}} = 2.10$, $M_{\text{human}} = 4.29$, $SD_{\text{human}} = 1.70$, $t = -0.97$, $d = -0.30$, $p = 0.34$ for awkward; $M_{\text{robot}} = 4.14$, $SD_{\text{robot}} = 1.46$, $M_{\text{human}} = 4.90$, $SD_{\text{human}} = 1.89$, $t = -1.46$, $d = -0.45$, $p = 0.15$ for watched).

Additionally, those in the robot condition said the role-playing experience would be more awkward if the game

Measure	Robot	Human
Dossier A solve rate	$M = 1.00$ $SD = 0.00$	$M = 1.00$ $SD = 0.00$
Dossier B solve rate	$M = 0.67$ $SD = 0.48$	$M = 0.52$ $SD = 0.51$
Dossier C solve rate	$M = 0.00$ $SD = 0.00$	$M = 0.00$ $SD = 0.00$
Puzzle phases solved	$M = 4.57$ $SD = 1.40$	$M = 4.29$ $SD = 1.90$
Total hints asked	$M = 5.52$ $SD = 2.69$	$M = 4.81$ $SD = 2.54$

TABLE I: Objective measures on puzzle performance and player engagement with the game guide. None of these measures demonstrated statistically significant differences between the Robot and Human conditions.

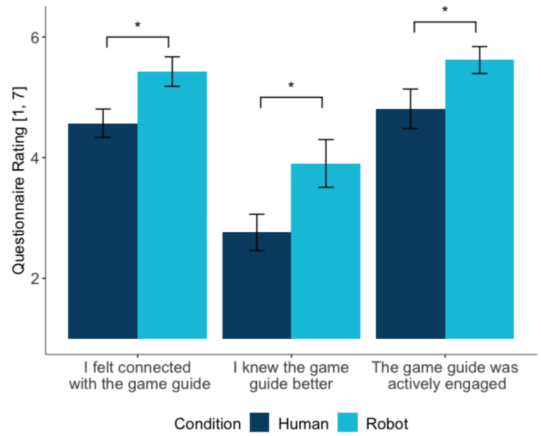

Fig. 6: Participants felt they were more connected with the game guide and that the game guide was more actively engaged as a robot than a human. (*) denotes $p < 0.05$. Error bars depict one standard error from the mean.

guide was a human ($M = 4.67$, $SD = 1.43$, $t = 2.14$, $d = 0.47$, $p = 0.045$). This differed from an expected baseline value of 4 (“neither agree nor disagree”) using a single sample t test, supporting H_2.

Furthermore, when asked about their overall impressions of the game guide, participants in both conditions described the game guide as helpful, understanding, and supportive. However, participants in the human condition also said the human actor was “*inorganic*” and “*overly formal*,” which caused them to feel awkward or “*be judged for asking for too much help*.” In addition, one participant in the human condition said the experience was “*a little weird*” due to it being “*a kind of interaction [they’d] usually have with an automated character in a video game*,” suggesting how “game master” roles may be more fitting for a robot.

When participants playing with a human game guide were asked how they would feel if the guide were replaced with a robot, they said they would be “*more comfortable asking for help*,” since a human may be “*secretly judging [their] lack of thinking*.” These qualitative observations show some anecdotal support for H_2.

C. Puzzle Performance

Table I summarizes objective measures on puzzle performance and player engagement with the game guide (e.g. asking for hints). All participants completed Dossier A, and none completed Dossier C. There is a trend of more participants solving Dossier B when playing with a robot game guide (14 of 21 participants) compared to a human (11 of 21 participants). On average, those in the robot condition also completed a greater total number of puzzle phases across Dossiers B and C. Additionally, they received a greater number of puzzle hints from the game guide, either by explicitly asking for help or when they tried to confirm their progress with the game guide. Nonetheless, we did not find statistically significant differences between conditions on these measures.

When those in the robot condition were asked how their experience might change if the game guide was human, one participant speculated that they “*would have been inclined to ask fewer questions*” and would “*take longer to solve the puzzles as a result*.” Another participant in the human condition said they would be “*less reluctant to ask [a] robot for help, since it feels like [they are] still working through the puzzle on [their] own*.” Those playing with the robot also “*enjoyed asking Agent Lee questions to see what his responses would be*,” which suggests they are likely to ask for more help. Taken together with participant performance data, we only find weak support for H_3.

D. Game Guide Social Attributes

Participants rated their perceived relationship with the game guide (Figure 6). They felt more connected with the robot ($M = 5.43$, $SD = 1.12$) than with the human game guide ($M = 4.57$, $SD = 1.08$, $t = 2.53$, $d = 0.78$, $p = 0.016$). Participants also thought they knew the robot game guide better ($M = 3.90$, $SD = 1.81$) than its human counterpart ($M = 2.76$, $SD = 1.37$, $t = 2.30$, $d = 0.71$, $p = 0.027$) after playing puzzles with them, and they believed that the robot game guide was more actively engaged in the game ($M = 5.62$, $SD = 1.02$) compared to the human actor ($M = 4.81$, $SD = 1.50$, $t = 2.04$, $d = 0.63$, $p = 0.049$). These ratings of participants’ relationship with the game guide support the idea that narrative robots facilitate stronger relationships compared to humans in the same role, showing support for H_4.

In Figure 7, participants rated the social attributes of the game guide relevant to H_5. While participants rated the sincerity of the robot game guide ($M = 5.01, SD = 1.27$) from the MDMT [32] as higher than the human guide ($M = 4.44, SD = 1.29$), this difference was not significant ($t = 1.45$, $d = 0.45$, $p = 0.16$). Additionally, they rated the game guide’s warmth using the RoSAS scale [31] as higher with the robot game guide ($M = 3.61, SD = 0.95$) compared to the human ($M = 3.06, SD = 0.89$), though this difference was not significant ($t = 1.95$, $d = 0.60$, $p = 0.058$).

Qualitative responses from participants support the idea that the game guide’s sincerity, warmth, and social presence

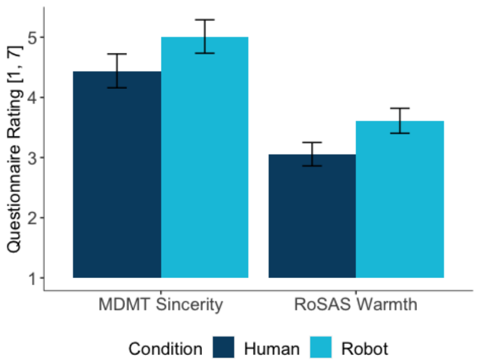

Fig. 7: Participants generally felt the game guide was warmer and more sincere as a robot, but without statistical significance. Error bars depict one standard error from the mean.

are magnified when they are a robot compared to a human. Participants in the robot condition stated that the game guide "*was more involved than [they] expected, interjecting with tips throughout which made him more than a physical manifestation of a hint button.*" They also "*felt like [they] could trust him for help*" due to the robot's facial expressions, and the game guide was "*very supportive in giving words of affirmation.*" In contrast, participants with the human game guide often described the guide as robotic and "*very synthetic*," such that they believed game guide's "*responses and encouragements [were] not genuine*." They also "*didn't connect like [they] might expect to with another person*," as the game guide felt "*one dimensional*." When considering these observations alongside participant ratings, we find that H_4 is supported by both statistical evidence and participant perceptions. However, we find only weak support for H_5.

E. Additional Observations

Responses to interview and survey questions indicate that most participants thought the game guide was helpful throughout the game. They were also immersed in the experience in both conditions, despite participants in the human condition expressing that having a robot game guide instead might require greater suspension of disbelief for the story to flow. Finally, many participants in both conditions said they wished the game guide would be more proactive, recognizing when they were struggling instead of giving hints only when asked by the participants. For some participants, the game guide did not contain a "*very strong presence*" in either condition, where they were "*so immersed in the puzzle*" that the game guide primarily acted as a voice that could aid the participant. However, participants noted that more time interacting with the game guide before beginning the dossiers would be helpful to "*build trust*."

V. Discussion

Our results show that participants have more fun, feel more connected with the game guide, and feel less judged when they play immersive puzzle games with a robot game guide compared to a human actor, which validates using robots as an entertainment medium as suggested by [34]. The perception of the robot game guide being more actively engaged and the increased interpersonal connection with the robot are also consistent with prior literature on games and entertainment with robots [20], [21], which suggest that emergent narratives mediated through a robot compared to a human have similar benefits to robots telling traditional narratives that do not require player interaction. Our results therefore support the idea that autonomous social robots can be effective alternatives to humans diegetically interacting with spectators in storytelling environments like escape rooms, theme parks, and interactive theaters because robots are able to create greater enjoyment compared to human actors, and the short-term interactions in entertainment spaces mirror those in our study.

Participants also felt less judged by the robot than the human game guide and indicated in their qualitative responses that they may be more willing to ask for help from a robot than the human game guide. These findings provide further support to the observation made by Leyzberg et al. that students are more willing to make mistakes and experience less social judgment from a robot compared to a human teacher [4]. Robots in a helper role that lower social judgment may then cause people to be more willing to ask for help, thus progressing the story or solving problems more quickly in themed environments.

Finally, the social presence and perceived increased engagement of a robot over a human in a primarily-passive role integrated into the environment supports the introduction of robots to unfamiliar public spaces where they need to be approached. For instance, people may be more likely to interact with and ask for help from a robot compared to a human in places like supermarkets, city streets, or airports if a robot is more attention-catching or seen as more emotive. People may also have a more positive outlook on a group of robots completing a task compared to a group of humans.

Limitations of our study include the choice of a tabletop Vector robot as the robot game guide, which may lead to a smaller social presence and less social pressure because of the robot's size compared to a human actor [9]. There was also a 1-2 second delay in the robot's responses to participants compared to the human actor. This delayed response speed may have an impact on participants' comfort and desire to interact with the robot [35], though players preferred the robot over the human game guide despite the delay. There is also a potential novelty effect associated with playing with a robot [36]. However, novelty is inherent in themed storytelling environments when raising people's enjoyment [37]. Thus, the higher ratings of enjoyment that were influenced by the novelty effect in the robot condition would still hold in real-world applications.

Overall, our results indicate that robot characters can be successfully integrated into a narrative game's storyline in place of a human, with minimal suspension of disbelief required. This may lead to a new direction in entertainment settings where short and simplistic yet fun and personalized interactions can be created by incorporating an embodied robot at a low cost compared to a human actor.

VI. Conclusion

We examined the differences between a robot and a human game guide in an immersive puzzle game where players must actively ask for help from the game guide to succeed. The game guide also played the role of both a cooperative partner and a "game master" within the game's narrative, emulating real-world entertainment spaces (e.g. escape rooms, theme parks) where robotic agents may be expected to have multiple social roles while portraying the same character. We find that human participants view a robot game guide as less judgmental, more actively engaged in the game, more connected with the player, and more fun to play with than a human game guide, suggesting that robotic systems could be implemented with great effect in interactive storytelling contexts to create scalable entertainment experiences.

References

[1] R. Brigante and S. A. S. Elger, "2020 Immersive Entertainment Industry Annual Report," 2020.

[2] H. Wax, "Engineering the Magic," *IEEE Women in Engineering Magazine*, vol. 2, no. 1, pp. 20–23, 2008.

[3] M. Wiemker, E. Elumir, and A. Clare, "Escape Room Games," *Game Based Learning*, vol. 55, pp. 55–75, 2015.

[4] D. Leyzberg, A. Ramachandran, and B. Scassellati, "The Effect of Personalization in Longer-Term Robot Tutoring," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 7, no. 3, pp. 1–19, 2018.

[5] C. A. Baker, "Creative choices and fan practices in the transformation of theme park space," *Transformative Works and Cultures*, vol. 22, 2016.

[6] M. Montola, "A ludological view on the pervasive mixed-reality game research paradigm," *Personal and Ubiquitous Computing*, vol. 15, no. 1, pp. 3–12, 2011.

[7] F. G. Pratticò and F. Lamberti, "Mixed-Reality Robotic Games: Design Guidelines for Effective Entertainment With Consumer Robots," *IEEE Consumer Electronics Magazine*, vol. 10, no. 1, pp. 6–16, 2021.

[8] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The Physical Presence of a Robot Tutor Increases Cognitive Learning Gains," in *Proc. of the Annual Meeting of the Cognitive Science Society*, vol. 34, no. 34, 2012.

[9] C. D. Kidd, "Sociable Robots: The Role of Presence and Task in Human-Robot Interaction," Ph.D. dissertation, Massachusetts Institute of Technology, 2003.

[10] F. Correia, S. Mascarenhas, R. Prada, F. S. Melo, and A. Paiva, "Group-based Emotions in Teams of Humans and Robots," in *2018 13th ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2018, pp. 261–269.

[11] G. Skantze, M. Johansson, and J. Beskow, "Exploring Turn-taking Cues in Multi-party Human-robot Discussions about Objects," in *Proc. of the 2015 ACM on Int. Conf. on Multimodal Interaction*, 2015, pp. 67–74.

[12] S. S. Sebo, P. Krishnamurthi, and B. Scassellati, "'I Don't Believe You': Investigating the Effects of Robot Trust Violation and Repair," in *2019 14th ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2019, pp. 57–65.

[13] N. Salomons, M. van der Linden, S. S. Sebo, and B. Scassellati, "Humans Conform to Robots: Disambiguating Trust, Truth, and Conformity," in *2018 13th ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2018, pp. 187–195.

[14] F. Correia, P. Alves-Oliveira, T. Ribeiro, F. S. Melo, and A. Paiva, "A Social Robot as a Card Game Player," in *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2017.

[15] J. E. Muñoz and K. Dautenhahn, "Robo Ludens: A Game Design Taxonomy for Multiplayer Games Using Socially Interactive Robots," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 4, pp. 1–28, 2021.

[16] P. Ralph and K. Monu, "Toward a Unified Theory of Digital Games," *The Computer Games Journal*, vol. 4, no. 1, pp. 81–100, 2015.

[17] H. Qin, P.-L. Patrick Rau, and G. Salvendy, "Measuring Player Immersion in the Computer Game Narrative," *Intl. Journal of Human–Computer Interaction*, vol. 25, no. 2, pp. 107–133, 2009.

[18] J. Kory and C. Breazeal, "Storytelling with robots: Learning companions for preschool children's language development," in *The 23rd IEEE Int. Symposium on Robot and Human Interactive Communication*, 2014, pp. 643–648.

[19] M. E. Ligthart, M. A. Neerincx, and K. V. Hindriks, "Design Patterns for an Interactive Storytelling Robot to Support Children's Engagement and Agency," in *Proc. of the 2020 ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2020, pp. 409–418.

[20] K. Hayashi, T. Kanda, T. Miyashita, H. Ishiguro, and N. Hagita, "Robot Manzai: Robots' Conversation as a Passive-Social Medium," *Int. Journal of Humanoid Robotics*, vol. 5, no. 1, pp. 67–86, 2008.

[21] I. Leite, G. Castellano, A. Pereira, C. Martinho, and A. Paiva, "Modelling empathic behaviour in a robotic game companion for children: an ethnographic study in real-world settings," in *Proc. of the 7th Annual ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2012, pp. 367–374.

[22] D. G. Bryant, J. Xu, Y.-P. Chen, and A. Howard, "The Effect of Robot vs. Human Corrective Feedback on Children's Intrinsic Motivation," in *2019 14th ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2019, pp. 638–639.

[23] A. Pereira, R. Prada, and A. Paiva, "Improving social presence in human-agent interaction," in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, 2014, pp. 1449–1458.

[24] N. C. Krämer, A. v. d. Pütten, and S. Eimler, "Human-Agent and Human-Robot Interaction Theory: Similarities to and Differences from Human-Human Interaction," in *Human-Computer Interaction: The Agency Perspective*. Springer, 2012, pp. 215–240.

[25] A. Weiss, A. Pillinger, K. Spiel, and S. Zauchner-Studnicka, "Inconsequential Appearances: An Analysis of Anthropomorphic Language in Voice Assistant Forums," in *Extended Abstracts of the 2020 CHI Conf. on Human Factors in Computing Systems*, 2020, pp. 1–7.

[26] H. Claure and M. Jung, "Fairness Considerations for Enhanced Team Collaboration," in *Companion of the 2021 ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2021, pp. 598–600.

[27] N. Tsoi, J. Connolly, E. Adéníran, A. Hansen, K. T. Pineda, T. Adamson, S. Thompson, R. Ramnauth, M. Vázquez, and B. Scassellati, "Challenges Deploying Robots During a Pandemic: An Effort to Fight Social Isolation Among Children," in *Proc. of the 2021 ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2021, pp. 234–242.

[28] W. Bainbridge, J. Hart, E. Kim, and B. Scassellati, "The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents," *Int. Journal of Social Robotics*, vol. 3, pp. 41–52, 10 2011.

[29] H. A. Rosyid, A. Y. Pangestu, and M. I. Akbar, "Can Diegetic User Interface Improves Immersion in Role-Playing Games?" in *2021 7th Int. Conf. on Electrical, Electronics and Information Engineering (ICEEIE)*, 2021, pp. 200–204.

[30] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing Robot Tutors to Individuals' Learning Differences," in *Proc. of the 2014 ACM/IEEE Int. Conf. on Human-Robot Interaction*. Association for Computing Machinery, 2014, p. 423–430.

[31] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The Robotic Social Attributes Scale (RoSAS): Development and Validation," in *2017 12th ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2017, pp. 254–262.

[32] D. Ullman and B. F. Malle, "Measuring Gains and Losses in Human-Robot Trust: Evidence for Differentiable Components of Trust," in *2019 14th ACM/IEEE Int. Conf. on Human-Robot Interaction*. IEEE, 2019, pp. 618–619.

[33] S. D. Gosling, P. J. Rentfrow, and W. B. Swann Jr, "A very brief measure of the Big-Five personality domains," *Journal of Research in Personality*, vol. 37, no. 6, pp. 504–528, 2003.

[34] D. Johnson, R. Cuijpers, K. Pollmann, and A. Ven, van de, "Exploring the Entertainment Value of Playing Games with a Humanoid Robot," *Int. Journal of Social Robotics*, vol. 8, no. 2, pp. 247–269, Apr. 2016.

[35] W. Wang, Y. Chen, R. Li, and Y. Jia, "Learning and Comfort in Human–Robot Interaction: A Review," *Applied Sciences*, vol. 9, no. 23, 2019.

[36] R. Aylett, *Games Robots Play: Once More, with Feeling*. Cham: Springer Int. Publishing, 2016, pp. 289–302.

[37] C.-H. Chang, S. Shu, and B. King, "Novelty in Theme Park Physical Surroundings: An Application of the Stimulus–Organism–Response Paradigm," *Asia Pacific Journal of Tourism Research*, vol. 19, no. 6, pp. 680–699, 2014.