An AI-Driven Card Playing Robot: An Empirical Study on Communicative Style and Embodiment with Elderly Adults

Michael Banck
AI and Knowledge Systems
University of Würzburg, Germany
michael.banck@uni-wuerzburg.de

Elisabeth Ganal
Socially Interactive Agents
University of Würzburg, Germany
elisabeth.ganal@uni-wuerzburg.de

Hanna-Finja Weichert
Socially Interactive Agents
University of Würzburg, Germany
hanna-finja.weichert@stud-mail.uni-wuerzburg.de

Frank Puppe
AI and Knowledge Systems
University of Würzburg, Germany
frank.puppe@uni-wuerzburg.de

Birgit Lugrin
Socially Interactive Agents
University of Würzburg, Germany
birgit.lugrin@uni-wuerzburg.de

ABSTRACT

The global population is aging, presenting growing challenges, particularly as shortages in social-related professions exacerbate these issues. Older age and related circumstances, such as the end of employment, often lead to social isolation and reduced cognitive stimulation, highlighting the need for technology-driven alternatives. This paper explores the use of socially interactive agents (SIAs) to address some of these challenges. Specifically, we implemented a prototype featuring either a physical or virtual embodiment and social- or task-oriented communication, which can autonomously recognize physical cards, game context and instruct moves. One focal point is the implementation, giving a broad overview of the architecture and important features. Equally relevant is the comparison of communication styles and embodiments, both known to influence users' perceptions when interacting with SIAs. We investigated the communication styles and embodiments to evaluate their influence on social presence, enjoyment, warmth, and competence in a user study with older adults. Our results revealed a positive effect on social presence with a physically embodied SIA in use and higher perceived competence for the social-oriented communication style. Furthermore, our findings confirm the proof-of-concept for the card detection AI and the gameplay strategy. This results in a promising tool for entertainment and cognitive stimulation and rewards us with overall positive feedback.

KEYWORDS

Card Game, Object Detection, Socially Interactive Agents, Social Robot, Interaction Study, Older Adults

ACM Reference Format:
Michael Banck, Elisabeth Ganal, Hanna-Finja Weichert, Frank Puppe, and Birgit Lugrin. 2025. An AI-Driven Card Playing Robot: An Empirical Study on Communicative Style and Embodiment with Elderly Adults. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 10 pages.

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

1 INTRODUCTION

Observing global demographic trends, it seems certain that growing challenges lie ahead. Both, the total population and the proportion of retired people relative to the employed are rising. As a result, the percentage of individuals aged 65 and older is projected to increase from 10% to 16% of the global population by 2050 [53]. The risk of age-related diseases also escalates with advancing age. Forecasts estimate that 66 million people aged 65 and older will be diagnosed with dementia by 2030, with the number expected to soar to 115 million by 2050 [47]. Dementia can cause declining memory, mood, and personality changes often leaving affected individuals unable to manage common tasks [47]. Elderly people, especially those with dementia, benefit from social interactions and cognitive stimulation, and social games, such as card games, are possibly useful to slow down cognitive decline and maintain brain functions [39]. Apathy is common in dementia patients, making it difficult for them to initiate social interactions on their own, while family members can struggle to find the time or means to motivate them [52]. Consequently, as the number of elderly people increases, the demand for caregivers and related professions will rise too. Given that social care roles, such as nursing, are already facing a shortage, this trend escalates the problem even further [50]. Facing these challenges, it becomes more and more relevant to explore alternatives.

The field of socially interactive agents (SIAs) offers a wide range of technical solutions to address not exclusively the aforementioned issues. These solutions include physical present social robots and intelligent virtual agents, displayed on screens, to enable socially interactive communication [43]. When leveraged with appropriate interactions and context, SIAs may help stimulate cognitive functions and reduce cognitive decline [8, 30]. Additionally, they can lower perceived social isolation through their ability to engage in social interactions [30, 31]. Some agents, e.g., the Paro robot, designed as a fluffy seal-baby for social assistance, aiming to reduce loneliness, can be employed with relative ease due to their simplicity [30]. Conversely, more complex systems often require supervision and intervention or an adjusted environment to function properly. These agents are engineered to perform sophisticated tasks, such as playing games or having meaningful conversations, but lack many times the capabilities to function autonomously [9, 10, 34].

We present a more complex system capable of playing a card game with elderly adults, alongside a user study, providing the engagement and time that family members or caregivers may lack. Furthermore, the SIA engages by actively playing, providing commentary, and instructing game-moves to increase cognitive stimulation and evoke emotional responses. First, the implementation is illustrated, including algorithms, AI, and interaction components, to design complex SIAs capable of performing advanced tasks. This approach reduces the common need for researchers to control the system manually, opening the door for an autonomous and scalable game setting. Second, we evaluate different SIAs to determine which communication style and embodiment is best suited on the dimensions social presence, enjoyment, warmth, and competence, when interacting with elderly people during a card game.

2 RELATED WORK

SIAs are virtually or physically embodied agents capable of autonomously communicating with people and each other in a socially intelligent manner using multimodal behaviors [44]. Introducing SIAs can improve many aspects, such as enhancing quality of life and reducing loneliness, especially for individuals with specific needs [2, 12, 28, 41]. In this section, we explore the usage of different communication styles for social agents, the influence of embodiment, and the deployment of SIAs in community games.

2.1 Social Agent Communication

Effective communication is key to successful interactions between humans and agents [19, 36, 49]. Considering this, some studies include task-oriented interactions, where communication is tailored to complete tasks [14, 51]. However, ideal communication should not only focus on task completion. User satisfaction goes beyond just getting the task done, and the most suitable communication style depends on the context and the type of interaction [24]. A task-oriented communication style could be used for formal, goal-oriented, or efficient interaction scenarios [54]. Zhu et al. [58] demonstrated that user preferences for communication styles vary depending on the environment, with a shift of competence towards task-oriented interactions in highly social crowding environments and vice versa for social-oriented ones. In an experiment by Taylor and Francis [15] using a chatbot, participants rated the social-oriented version higher for social presence and enjoyment when compared to the task-oriented one. However, no effect on trust and intention of use was found. Van Dolen et al. [54] examined the effect of task- versus social-oriented communication in an online commercial group chat. The advisor in this chat used either the task or social communication style. While task-oriented was rated higher for perceived speed of delivery and reliability, the social-oriented style was superior in enjoyment for session satisfaction. However, Ganal et al. [21] reported even in the field of information management, where efficiency is a main goal, a social-oriented chatbot is favored over a task-oriented one. The interaction with the social-oriented chatbot led to a significantly higher rating for service quality, enjoyment, and interpersonal competence when compared to the task-oriented bot. Similarly, Verhagen et al. [55] investigated how expertise in virtual customer service agents influences social presence differently in a social-oriented communication style compared to a task-oriented approach. Generally, card games are usually played with a certain pace, where game moves are not expected to take too long. At the same time, the opponent should demonstrate competence and friendliness to support a rewarding gameplay experience. These elements could favor either social- or task-oriented communication.

2.2 Physical and Virtual Agent Embodiments

SIAs can be either physically or virtually embodied, each with distinct characteristics and individual (dis-)advantages in terms of e.g., scalability, social presence, or animation capability [43]. Comparing physical SIAs to virtual screen-based SIAs is essential to find the key features that influence user experience and performance, especially in community game settings. Leite et al. [38] conducted a study using chess as a game. Participants played against either a physical iCat robot or a virtual 3D model of iCat displayed on a screen. The results indicate greater enjoyment when playing against the physical iCat. Similarly, Wainer et al. [56] investigated watchfulness, enjoyment, and helpfulness by designing an experiment based on solving the tower of Hanoi puzzle. They evaluated three different settings: a physical robot present in the same room, a video of the physical robot, and a virtual robot. Participants rated the physical robot in the same room as more helpful, enjoyable, and watchful. However, they could not find significant differences in task performance across the three settings. Additionally, no notable preferences between the virtual robot and the video of the robot could be concluded. Using virtual and physical embodied SIAs for physical exercises with older people, the physical robots were again rated as more enjoyable and useful [16]. Furthermore, studies have demonstrated not only the superiority of physically present robots, in many cases, for aforementioned enjoyment, but also for reliability, trustworthiness, credibility, and social presence [32, 33, 40]. Nevertheless, some studies indicate that virtual SIAs perform comparably in certain aspects and can influence user perception similarly to their physically present counterparts, making the choice of embodiment also context-dependent [3, 27].

2.3 Computer Aided Card Games

The integration of computer systems into social games and therapies has become increasingly prevalent, driven by advancements in AI and human-computer interaction. These systems aim to enhance traditional approaches by introducing new forms of interaction, engagement, and accessibility, particularly for people who benefit from augmented social experiences, such as the elderly or those with limited mobility [45]. In the last decade, several studies explored computer-aided technologies to facilitate social games. Physical and virtual SIAs have been used in games for various purposes, serving different social roles. They help create and sustain the social dimensions of games, improving social immersion and providing a better game experience for players [46]. Correia et al. [10, 11] showcased the Portuguese card game Sueca. Utilizing a table with integrated multitouch screen and cameras, the authors modeled a playable interface. This involves designing new Sueca cards, including QR-Codes to allow identification by the computer vision framework *reacTIVision*, while keeping the regular symbols for human players [9, 29]. The participants used these cards, while the agent scanned the QR codes of its cards at the beginning and played the cards virtually on the screen. To embody the SIA, an

EMYS robot was used, while to control social interaction, the agent utilized game events and states representing emotions. However, one issue the participants pointed out was the card design with the QR codes being confusing, and on few occasions, cards were mistaken [9]. Similarly, Fischbach et al. [17] integrated a Reeti robot with a multitouch table in a role-playing game, to potentially improve game immersion. Kim and Suzuki [34] explored a SIA playing Texas Hold'em poker. While the cards remained unaltered, the movement and speech of the robot were controlled using Wizard-of-Oz techniques. Additionally, a professional dealer was part of the setting, assisting the robot by handling the cards and moving the chips. Although the system yielded interesting findings in human-robot interaction, it fell short in terms of autonomy. In contrast, the assistive robot Brian 2.0 and its later version, Brian 2.1, demonstrated high-level, computer-determined interactions, although they did not actively play games [5, 42]. Instead, the participants played a memory card game, while the robot provided support e.g. by suggesting matching cards when players could not recall them. To this end, an entire Human-Robot interaction (HRI) control architecture was developed to evaluate the participant's state. This setup included a camera, microphone, and heart rate sensor to determine the affective arousal level. Combined with other internal states and reinforcement learning, Brian 2.1 provided verbal and non-verbal cues to support elderly participants in playing the game and maintaining their motivation [42].

In conclusion, gathering all the necessary information to physically play a card game and incorporating social interaction using SIAs remains challenging. Although Wizard-of-Oz methods effectively address these challenges and cover most research questions, our motivation is to develop a system that can investigate our objectives with autonomous SIAs, requiring less support from researchers and fewer alterations of the game elements. Therefore, this paper presents the implementation and evaluation of a prototype designed to play cards with elderly people. It can switch between task- and social-oriented communication styles. Additionally, the embodiment of the SIA is interchangeable, allowing for either a physical or virtual embodiment. The goal is to determine which communication style and embodiment is most effective for playing cards with elderly individuals. To evaluate this, we make use of the dimensions: social presence, enjoyment, warmth, and competence. To this end, we define five hypotheses:

H1: SIAs using the social-oriented communication style evoke a greater sense of social presence compared to the task-oriented one.

H2: Playing with a physically embodied SIA leads to a greater sense of social presence compared to the virtually embodied SIA.

H3: Playing with a physically embodied SIA results in more enjoyment compared to the virtually embodied SIA.

H4: A SIA using the social-oriented communication style leads to a greater sense of warmth compared to the task-oriented one.

H5: A SIA using different communication styles affects the perceived competence.

3 SYSTEM OVERVIEW

This section presents the overall system, introducing important game elements and software components that make up the prototype. It lays the foundation for understanding how the game progresses from the game start over alternating turns to the game end (game flow), as well as how the SIA calculates game-moves (gameplay strategy) and the interplay of game elements and users. Fig. 1 shows physical elements and the simplified interaction of the software modules described in this section.

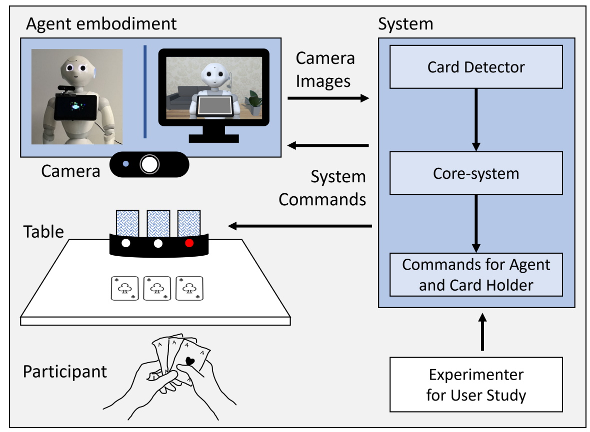

Figure 1: System overview and study setup for user study.

3.1 Card Game

Selecting an appropriate card game is crucial, as several factors determine the suitability, with simplicity being a key consideration. The game should be easy to learn or, ideally, well-known. Additionally, a balance of luck and skill is important to keep it engaging and enjoyable for both beginners and experienced players, preventing it from becoming predictable too quickly. *Schnauz* (also called Schwimmen) is such a game and is widely played in Bavaria, Germany. Schnauz is similar to *Thirty-One*, which is played in the USA and the UK. This game is ideal due to its straightforward rules and widespread recognition. Breaking down the rules, each player receives three cards at the beginning, and three additional community cards are placed face up in the center of the table. The game proceeds in turns where each player must perform one of four possible game-moves:

(1) Skip - do nothing, skip a turn
(2) Knock - announces the end of the game, which will occur in the next round
(3) Swap one card - exchange one card from the hand with one form the community cards.
(4) Swap all cards - exchange all hand cards with all community cards.

The goal is to maximize the score of the three cards in hand by collecting cards of the same suit or rank. Typically, a Bavarian card deck is used. It consists of four suits (Hearts, Bells, Acorns, and Leaves) and nine ranks (6 to 10, Under Knave, Over Knave, King, and Ace), similar to the French deck.

3.2 Physical Components

The setup includes key components critical for environment recognition, communication, and representation to ensure a seamless game flow.

Camera: The Logitech Brio Stream is a 4K camera and captures the game in real-time. For the physical embodied version, the camera is mounted on Pepper's tablet. For the virtual SIA, it is placed on top of the monitor. In both settings, the camera is approximately 40 cm above the table and 50 cm in vertical distance to the community cards. This ensures a clear view of the cards, minimizes distortion, and maintains image quality. The captured images are transmitted to the PC for further processing by the object detection module.

Card-Holder and Table: The custom-made card-holder includes an Arduino connected to the laptop and an LED for each card slot. It organizes the SIA's cards so that the camera can capture them while only the backs are visible to the opponent. The core-system sends messages to the Arduino via a serial connection before and after the SIA's card swap. The swap is indicated by blinking the corresponding LED. Due to Pepper's height, we used a lower table (45 cm) to ensure a good camera angle and to restrict the area where the cards could be played, while still offering participants a clear view and keeping all elements within their reach.

PC: Computationally intensive tasks, such as AI operations, are managed by a high-end laptop. This PC is equipped with an Intel Core i7-13700H CPU, 16 GB RAM, and a NVIDIA GeForce RTX 4070 with 8 GB VRAM. Consequently, this setup allows real-time processing of the camera feed and executing the gameplay strategy and game flow simultaneously. In addition, the PC also serves as the central hub, managing communication between the card-holder, the camera and the physical or virtual SIA.

Pepper Robot: Pepper serves as the physical embodiment in our system, positioned on one side of the table, facing the seated user at eye level. It is equipped with various sensors and actuators, including speakers, microphones, and more. In this study, only the gestures and Text-to-Speech (TTS) capabilities are utilized, while the other functions are deactivated or unused. Pepper's behavior is controlled via the NAOqi SDK version 2.5.6 for Java, providing verbal instructions, commentary, and gestures during the game.

Virtual Pepper: To ensure the virtual SIA closely mimics the physical Pepper robot, its appearance and movements were replicated using the Pepper Python Unity Toolkit (PePUT) [20] and Unity version 2019.4. PePUT provides a virtual Pepper robot, it can translate the regular Pepper behaviors into Unity movements and map them to the virtual skeleton. The virtual robot is placed close to the virtual camera. The scene includes a hardwood floor, wallpaper, a couch, a lamp, a bookshelf, and a plant for atmosphere. The virtual Pepper performs the same actions in response to the same events as the physical SIA. This is illustrated in Fig. 1. Similarly to the appearance, the voice was copied, too. All comments were generated beforehand. The core-system coordinates the replay of the correct audio files with the behaviors.

Monitor: The virtual SIA embodiment is a digital representation of the Pepper robot. In contrast to the physical robot, the virtual robot is displayed on a full-HD 24-inch monitor. Similar to Pepper, the monitor is placed in front of the table at an elevated position, ensuring optimal visibility for the user.

3.3 Software Components

The software architecture is designed to handle user interaction, real-time camera input, object detection, and the game flow. It is implemented through various modules and threads that comprise the technology stack of the project.

Card-Detector: The Card-Detector module is a Python program that continuously captures the camera stream. These frames are processed and a retrained YoloV10 [57] model is used to detect the cards and their location. With these locations, the cards are clustered into community, player and SIA blocks. Once a frame is processed, the data is sent to the core-system through an HTTP request. Additionally, the program can display the currently captured image, including card detections and clusters in a separate window. This feature is for the experimenter only, to ensure that all cards are within the camera's view and correctly detected.

Core-System: The core-system as a hub of the entire system, needs to orchestrate all other modules and make final decisions, this requires parallel processing and synchronization. Considering this, Kotlin was used due to its capabilities in multi-threading environments. Upon startup, it initializes all other modules and handles their shutdown and termination. The gameplay strategy and flow are managed by a finite state machine, with the different phases and transitions providing events for actions used by the SIA. For example, at the beginning of a turn, the SIA can say, "It is your turn" and make a gestural indication such as pointing at the player. All gestures and the TTS are achieved by leveraging the NAOqi SDK version 2.5.6 and a RESTful API created in Unity, enabling data to be sent in a similar manner, resulting in identical behaviors across both physical and virtual embodiments.

Experimenter Interface: An Angular application presents the current system state and provides controls to the experimenter. It ensures correct system operations and allows for intervention when participants want to *knock* or *skip*. When participants vocalize these actions, the experimenter acts as a speech recognition AI and presses the associated buttons.

4 IMPLEMENTATION

This section describes the system implementation, emphasizing key software components designed to work seamlessly with both physical and virtual embodiments. It highlights algorithms, programming frameworks, and details that foster a cohesive and interactive gameplay experience. [1]

4.1 Card Detection

A fundamental aspect of the gameplay with physical cards is the ability to recognize them correctly using computer vision. This subsection introduces the pipeline and steps, used to perform the card detection and provides results for the proof-of-concept.

Card Detection Architecture: The card detection starts with the webcam capturing a continuous video stream processed via OpenCV [13]. After capturing and resizing, the frame is forwarded to the YOLOv10 framework. Initially introduced in 2016 and refined through subsequent iterations [26, 48, 57], YOLO (You Only Look Once) is optimized for real-time inference and has been adapted for various domains, including autonomous driving and medicine [35, 59]. Given the absence of pre-labeled datasets for the specific cards in use, a YOLOv10-X model was retrained using a Bavarian card deck with larger symbols for seniors. To bridge the gap

[1] Details on the system implementation are shared for scientific purposes upon request.

in labeled data, a synthetic, photo-realistic dataset was generated using Blender version 3.6.1. For this purpose two custom plugins were developed: The first plugin randomizes card positions, combinations, lighting, tabletops, and camera angles to generate many distinct game situations, while the second plugin extracts annotations and outputs bounding box labels. Using this process 25,000 images for training and 5,000 for validation were created. As a result, the retrained model was able to locate and distinguish the 36 different cards of the game. As a post-process, all predictions including duplicated labels were reduced in favor of the one with the highest confidence. Unlike other detection tasks, in most card games a single label can be present only once. Using this property, the detected cards are then sorted by their distance to the camera, assuming hand cards are closest and opponent cards are farthest. This sorting heuristic performs reliably when the opponent sits across the table. Based on the ordered cards, clusters are formed for the community cards and each player's hand cards. The clusters and card detections are showcased in Fig. 2. The big clusters labeled 1, 2, and 3 represent the hand, the community, and the opponent's cards, respectively.

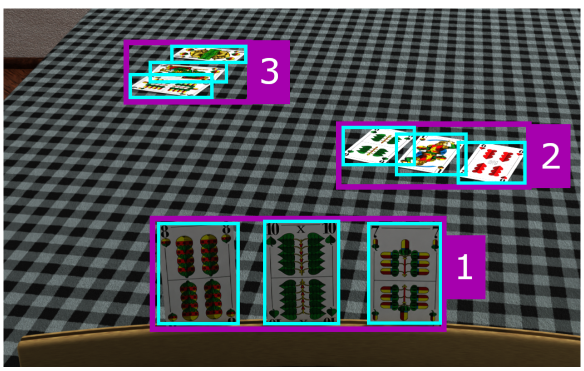

Figure 2: Card Detection and clusters in a Blender scene.

The clusters and their cards are forwarded from python to the core-system using an http request. To ensure stability, a stack of all predictions is maintained, and card swaps are only recognized if persistent over a predefined window size, mitigating the effects of transient occlusions. The window size tends to vary depending on the hardware specs, while faster hardware requires larger windows to ensure stability over the same period of time.

Proof-of-Concept: The training was based on synthetic images, but the performance needs to be confirmed on real images too. In the absence of publicly available data, we created a small dataset for a proof-of-concept. For this 252 images were captured, facing different angles, lighting conditions, and distances, resulting in 1260 labeled cards (35 for each of the 36 cards). These images were recorded using the same webcam used in our system to replicate realistic conditions. The evaluation was performed using retrained YOLOv10 models: *YOLOv10-X* and *YOLOv10-M* with image sizes of 640^2, 960^2, and 1080^2 pixels. *TP* denotes all predicted bounding boxes that match a ground truth bounding box with an Intersection over Union (IoU) of at least 50%. All models were trained with the same dataset and the same parameters, except for batch size adjustments to optimize the 24 GB VRAM usage. The training process included *early stopping* and *patience* set to 15, meaning models trained until no further improvement was observed for 15 epochs, resulting in different epoch counts for each model. Tab. 1 presents the results of our proof-of-concept dataset. All tested models produced promising results and illustrated generalization on the real image data. Larger models and input sizes need more memory and inference time. In our setting, where the detections serve as the single source of truth and as a control element of the game, making as few mistakes as possible is the most important factor. Therefore, the model of choice is the YOLOv10-X model with a 1080^2 input size.

Table 1: Performance of various YOLOv10 models and input sizes.

Model	Size	TP	FP	FN	Precision	Recall	F1 Score
X	1080	1260	0	0	1.000	1.000	1.000
M	1080	1246	2	12	0.998	0.991	0.994
X	960	1233	10	17	0.992	0.986	0.989
M	960	1231	10	19	0.992	0.985	0.988
X	640	1219	9	32	0.993	0.974	0.984
M	640	1131	37	92	0.968	0.925	0.946

Note. TP = True Positive, FP = False Positive, and FN = False Negative.

4.2 Gameplay Strategy

Schnauz like many other games, relies on both luck and strategy. In our experiments, the objective is to maintain a balance between challenging but not discouraging the opponent by implementing a superior strategy such as the Monte Carlo Tree Search [4]. Taking this into account, we designed and validated our gameplay strategy.

Gamplay Strategy Algorithm: In each turn, Schnauz offers 12 possible moves: skip c, knock k, swap all s_{all}, or swap one $s_{x,y}$ card, where x and y represent the index of the agent's hand card and the community card, respectively, resulting in 3^2 single-swap options. Given this small search space, all possible outcomes can be computed and evaluated by the gameplay strategy in a reasonable time. The resulting scores are based on the current hand S_c and all potential hands $S_{x,y}$ and S_{all} that can be achieved through $s_{x,y}$ or s_{all}. Based on the hand cards H and the community cards C, the game-move function gm is defined as follows:

$$\mathrm{gm}(H,C) = \begin{cases} k, & S_c = \max(S_c, S_{all}, S_{1,1}, \ldots, S_{3,3}) \text{ and } S_c > 21 \\ c, & S_c = \max(S_c, S_{all}, S_{1,1}, \ldots, S_{3,3}) \text{ and } S_c \leq 21 \\ s_{all}, & S_{all} = \max(S_c, S_{all}, S_{1,1}, \ldots, S_{3,3}) \\ s_{x,y}, & \text{otherwise} \end{cases} \quad (1)$$

As presented in equation 1, if the current hand score is optimal, it will either knock k or skip c, depending on whether the score exceeds 21. This is to avoid premature knocking and maintain the game flow. The score 21 is selected as threshold since it is the highest score of two matching cards, while every higher score requires all three cards to match. Although this strategy results in optimal game-moves given the immediate information, the agent does not track the opponent's cards. Therefore, swaps can change the community cards in favor of the opponent. Consequently, the agent remains competitively balanced to win some rounds based on skill rather than pure luck.

Gameplay Strategy Evaluation: The gameplay strategy is simple. However, testing it against a perfect or inferior strategy does not convey its intended balance. Instead, only its performance against human players can illustrate the win-loss ratio it is designed for. To this end, we conducted 50 rounds against the gameplay strategy, with the strategy losing 23 rounds, winning 23, and tying 4. Not considering ties, this results in an exact 1:1 win-loss ratio. While more rounds may reveal trends, this already confirms a balanced and suitable gameplay strategy for the study.

4.3 Game Flow Events

The SIAs social interaction is tied to the state machine and the study configuration. Schnauz, as a turn-based card game, was modeled including five states, illustrated in Fig. 3. The game starts in the *Before Game Started* state, transitions to the *Human Player* state, and oscillates between *Human Player* and *Bot* until a player knocks, moving to either *Round Over* or *Game Over* when the predefined number of rounds is exceeded. Each state triggers various events that control the SIA's behavior. Some events are conditional and rely on decisions the gameplay strategy has made, like: skip, knock, swapAll, and swapOne. In addition, there are events providing parameters. For example, swapOne has four parameters: *handCardName*, *communityCardName*, *handCardPosition*, *communityCardPosition*. These parameters allow dynamic TTS phrases to communicate contextually, such as: "Would you please swap my left card for the King of Hearts from the community cards?".

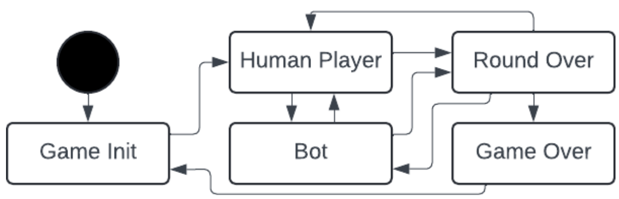

Figure 3: Schnauz Finite States

4.4 Communication Style Design

The game flow events described in section 4.3 are fixed, tied to the game flow, and cannot be altered other than through programming. However, the usage and content of these events can be modified using a configuration file. This game contains 56 events triggered throughout the game. Some events are invoked multiple times, while others are used only once, or depending on the individual game may not be triggered at all. Based on the study and its goals, a subset of these 56 events is used to design the SIA's communication. For each event, multiple alternative TTS sequences and behavior tuples can be defined. The tuples consist of two elements: the TTS part, including text and parameters such as volume, and the behavior part, which uses strings for predefined movements that resemble gestures. For the study, two different communication styles were explored: social-oriented and task-oriented, with the content based on existing concepts [7]. The social-oriented version utilized 44 of the 56 events, filling them with 189 sequences, resulting in an average of 4.02 alternative sequences per event. On the other hand, the task-oriented version used 29 events with a total of 77 sequences, averaging 2.66 sequences per event.

Table 2: Examples of Communication Style Content.

Event	Task-Oriented	Social-Oriented
win	You won with a score of #{0}.	Congratulations! You are the winner of this round with a score of #{0}.
think	-	Hmm, what should I do next...

Note. #{0} marks a placeholder

Tab. 2 showcases the events *win* and *think*. The *win* event is triggered when the round is over and the human player has a higher score than the SIA, while *think* is triggered at the beginning of the SIA's turn. The task-oriented style provides minimal communication, conveying only information, whereas the social-oriented style includes additional engaging, non-informative phrases, potentially leading to longer turns. The *win* event also uses placeholders (e.g., #{0}), which are replaced by the SIA at runtime. Nonverbal communication is another aspect of the SIA's interaction. These behaviors are synchronized with the TTS elements. However, behaviors are optional and not required for every TTS element. The task-oriented and social-oriented communication utilize the same behaviors. The social-oriented style subscribes to more events and includes more TTS-behavior tuples that actively use a behavior. Finally, Pepper's predefined behaviors matched all the required situations, eliminating the need for custom behaviors.

5 USER STUDY

The study was conducted using a 2×2 mixed design. This design was chosen, as the recruitment of elderly individuals of the right age who are also able to visit the facilities in their spare time is challenging. Further, playing with the Pepper robot became part of our promotional effort to generate more interest. To honor this promise, every participant should have the opportunity to interact with the physical robot. The within-subjects factor corresponds to the SIA embodiment to compare the physically embodied SIA with the virtual SIA. The between-subjects factor was defined by the two conditions, in particular the variation of the communication style of the SIA (social-oriented vs. task-oriented). The study was approved by the ethics committee of the Institute Human-Computer-Media at the University of Würzburg.

5.1 Measures

The *Almere* model [23] was developed to investigate the acceptance of SIAs among older people, of which we used the subscales *Social Presence* and *Enjoyment*. With *Social Presence* we investigate whether the physically embodied agent is perceived as more socially present than the virtual one. *Enjoyment* is assessed to compare which embodiment is perceived as more enjoyable when participants interact with the SIAs. Both subscales are measured with five items on a 5-Point Likert scale ranging from 1 "strongly disagree" to 5 "strongly agree", e.g., "When interacting with the robot I felt like I'm talking to a real person." (Social Presence) or "I enjoy doing things with the robot." (Enjoyment). Heerink et al. [23] reported a reliability of Cronbach's α = .83 for the subscale *Social Presence* and α = .76 for *Enjoyment*. The respective reliabilities for the current sample were α = .75 (Social Presence) and α = .82 (Enjoyment). To measure the perceived warmth and competence of the SIA, we used a questionnaire by Fiske et al. [18] with the subscales *Warmth* (e.g., "friendly") and *Competence* (e.g., "intelligent"). Both subscales include six items answered on a 5-point Likert scale ranging from 1 "strongly disagree" to 5 "strongly agree". Warmth and competence are reported with reliable Cronbach's α values in previous research [25]. The respective reliabilities in the current sample were Cronbach's α = .88 (Warmth) and α = .89 (Competence).

For general demographic data, participants provided their numeric age and gender (male, female, diverse). Further, they reported their previous experience with social robots and how often they play cards. In addition, participants were asked which embodiment of the robotic agent they preferred and how they liked the card game Schnauz.

5.2 Setup and Study Procedure

The study was conducted in one of our laboratories as well as in two retirement homes. First, the experimenter greeted the participants and explained the study procedure, including the interaction with the robot and the use of questionnaires. Afterwards, they were informed about the voluntariness of participation and anonymity of the data, data protection, and information on data collection. All participants gave their informed consent. The experimenter explained the rules of the card game and subsequently played at least three rounds with the participants to ensure proper understanding.

Then, participants were randomly assigned to one of the two conditions (social- vs. task-oriented) and played cards for five rounds each with the physical and the virtual robot in randomized order. The two card playing blocks are each followed by answering the questionnaire part about the play experience and the perception of the SIA. The last part of the questionnaire assessed the participants' experience with social robots, personal frequency of card playing, and demographic data. At the end of the study, the participants were debriefed, allowed to share feedback, ask questions, and thanked for their participation. The study took about 45–60 minutes.

5.3 Participants

Participation in our study required people to be 55 years or older. They were recruited via colleagues and acquaintances, flyers, retirement and residential homes, and local community centers. We recruited 70 people who took part in the study. However, $n = 5$ participants were excluded from the statistical analysis because they did not complete all questionnaires. Nevertheless, those participants remain in the technical analyses. In total, $N = 65$ participants (39 female, 25 male, 1 diverse) between 55 and 90 years ($M_{age} = 71.48, SD_{age} = 8.57$) took part in the study and completely filled in the questionnaires. The social-oriented and task-oriented SIA conditions comprise 33 and 32 participants, respectively. Among all participants, $n = 3$ had prior experience interacting with social robots. Further, $n = 10$ participants reported playing cards more than eight times a month, $n = 13$ between one to seven times a month, $n = 26$ several times a year, and $n = 16$ never. The card game *Schnauz* was already known by $n = 24$ participants.

5.4 Results

All statistical analyses were conducted using *JASP* [2] version 0.18.3.0 and a significance level of .05. The descriptive data for the questionnaire results regarding social presence, enjoyment, warmth, and competence can be found in Tab. 3. The assumption of normal distribution is given for the subscales *Social Presence* and *Warmth* based on the results of the Shapiro-Wilk test ($p > .05$). For the subscales *Enjoyment* and *Competence* the assumption of normal distribution is violated due to the Shapiro-Wilk test ($p < .05$). Levene tests were carried out to verify the homogeneity of variance, which was confirmed for each of the individual analyses ($p > .05$). As the ANOVA is robust to violation of the normality assumption [22], we carried out 2×2 mixed ANOVAs for the data analyses, with communication style as between-subject factor and embodiment as within-subject factor.

Table 3: Descriptive values for data on agent perception.

	Task-oriented				Social-oriented			
	physical		virtual		physical		virtual	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
SP[a]	2.66	0.88	2.34	0.87	2.82	0.81	2.73	0.88
ENJ[a]	3.85	0.80	3.83	0.67	4.09	0.87	4.07	0.90
W[a]	3.41	0.88	3.22	1.12	3.69	0.78	3.63	1.01
C[a]	3.87	0.94	3.70	1.09	4.19	0.64	4.25	0.74

Note. SP = Social Presence, ENJ = Enjoyment, W = Warmth, and C = Competence. [a] Calculated values from 1 to 5.

Social Presence: The results revealed no significant main effect of communication style, $F(1, 63) = 2.041$, $p = .158$, partial $\eta^2 = .031$. However, the analysis showed a significant main effect of embodiment, with $F(1, 63) = 5.053$, $p = .028$, partial $\eta^2 = .074$, confirming that participants reported higher levels of social presence when interacting with the physical compared to the virtual one. It showed no significant interaction between communication style and embodiment, $F(1, 63) = 1.563$, $p = .216$, partial $\eta^2 = .024$.

Enjoyment: The ANOVA for *Enjoyment* showed no significant main effects, neither for the communication style ($F(1, 63) = 1.615$, $p = .208$, partial $\eta^2 = .025$) nor the embodiment ($F(1, 63) = 0.071$, $p = .790$, partial $\eta^2 = .001$). Further, the analysis revealed no significant interaction between communication style and embodiment regarding *Enjoyment*, $F(1, 63) < .001$, $p = 0.997$, partial $\eta^2 < .001$.

Warmth: The results showed no significant main effects, neither for communication style ($F(1, 63) = 2.548$, $p = .115$, partial $\eta^2 = 0.039$) nor embodiment ($F(1, 63) = 1.848$, $p = .179$, partial $\eta^2 = .029$). Further, the analysis revealed no interaction between communication style and embodiment, $F(1, 63) = 0.483$, $p = .489$, partial $\eta^2 = .008$.

Competence: The data revealed a significant main effect in favor of the social-oriented communication style, $F(1, 63) = 4.685$, $p = .034$,

[2]https://jasp-stats.org/

partial η^2 = .069. However, the analysis showed no significance for the main effect of embodiment $F(1, 63) = 0.415$, $p = .522$, partial η^2 = .007 and no significant interaction effect $F(1, 63) = 1.908$, $p = .172$, partial η^2 = .029.

Embodiment Preferences: Among all participants, $n = 35$ preferred the physical embodiment, whereas $n = 23$ preferred the virtual one, and $n = 7$ voted for both.

Rating of the card game: Only $n = 3$ participants indicated that they did not like the card game *Schnauz*, whereas $n = 6$ rated it as average, $n = 32$ good, and $n = 23$ as very good, and one participant did not rate the card game.

Gameplay Strategy Evaluation: The 70 participants played against both the physical and virtual SIA. Excluding ties, which were considered balanced outcomes, participants recorded 71 wins and 56 losses in the remaining 127 of 140 games, resulting in a 56% win rate for participants.

6 DISCUSSION

In our study, we evaluated a prototype with a virtual and physical robotic embodiment that played a physical card game with elderly users in two conditions, either with task-oriented communication only, or enhanced by the social-oriented communication style. We did not find the expected effect of communication style on social presence, and thus reject H1. However, our results show a tendency in the expected direction, in line with findings by Chattaraman et al. [6], who reported higher ratings for the social-oriented interaction style. A possible reason for the absence of a significant effect might be that the interaction time was too short or that the communication style designs were not sufficiently distinguishable, particularly as the between-participants factor in a setting that is likely to provoke a novelty effect. Furthermore, in line with our expectations, we found a significant main effect of embodiment on social presence, suggesting that the physically embodied SIA had a positive effect, thus accepting H2. This finding reflects studies in other domains, using social robots that highlight the benefit of physically embodied agents [1, 37], and confirms it for the domain of card playing. In contrast to our expectations, we did not find a significant effect on enjoyment, leading to a rejection of H3. This contradicts prior research claiming that physical embodiment improves enjoyment [38, 56]. However, the overall ratings for enjoyment, presented in Tab. 3, in the experiment were high, with slightly higher ratings in the social-oriented versions. The high perceived enjoyment is also reflected in the card game's high ratings, suggesting participants enjoyed playing regardless of the embodiment. Thus, we assume a ceiling effect. Similarly, we did not observe significant differences regarding the perceived warmth of the SIA, resulting in a rejection of H4. This is somewhat surprising, considering that the social-oriented agent was specifically designed to increase the social factor and thus being perceived as more sympathetic. However, there is a small trend in the assumed direction, showing higher values for both embodiments in the social-oriented compared to the task-oriented version. As expected, the results showed a significant effect of communication style on perceived competence. Our data revealed an improvement when the SIAs demonstrated social interaction. This finding is particularly interesting, as it highlights the need for social interaction for a SIA in the domain of card playing. Since playing cards is a social activity per se, a purely task-oriented interaction might not have been perceived as competent. Similar to our findings, Ganal et al. [21] reported significantly higher interpersonal competence for a social-oriented chatbot compared to a task-oriented one.

Similar to the previous proof-of-concept, the card detection and the entire system performed well when used by the participants. Although no image data was collected during the study, the experimenter reported overall reliable card detection, with only minor imperfections observed in a few games due to poor lighting.

In sum, the system performed well with elderly users, while the data of the interaction study suggests that a physical body outperforms a virtual one in terms of social presence, but not in terms of enjoyment, which was generally high over the conditions. In addition, regardless of the embodiment, social behavior for SIAs is an important asset to be perceived as more competent.

7 CONCLUSION AND FUTURE WORK

Finding effective methods for interaction between elderly people and computers is becoming increasingly important due to the demographic change. The right concepts could slow down cognitive decline, reduce loneliness and provide many other positive effects, some of which were explored in our experiment. We demonstrated the successful implementation of a SIA prototype with two embodiments, customizable communication style and autonomous capabilities. A 2×2 mixed user study was conducted including 65 older adults, who played a card game against the physical and virtual embodied SIA, featuring either a social- or task-oriented communication style. The results revealed an increase in social presence when the SIA is physically embodied, and higher ratings of competence for the social-oriented communication style. No significant effects were observed for enjoyment, while the overall ratings were high, and participants liked the overall experiment. The missing effect on warmth will be explored in future studies with the more promising physically embodied SIA. To this end, we investigate different styles, based on social-oriented communication, such as egoistic vs. altruistic communication. Furthermore, communication could also be enhanced by extending the system so that interactions are not exclusively triggered by game-flow events but also by integrating small talk on topics unrelated to the game.

Additionally, the card detection and gameplay strategy were evaluated, and the results confirmed the proof-of-concept for these components, while also demonstrating reliability during the user study. Future work could explore multiplayer setups, as some participants mentioned in their feedback card games designed for more than two players.

Ultimately, receiving mainly positive feedback and generating interest motivates us to continue exploring our approach further. We see our prototype as a feasible initial step that holds potential to improve the quality of life for elderly individuals.

ACKNOWLEDGMENTS

This research was partly funded by the European Social Fund Plus (ESF+) under Project No. 1-2-230972.

REFERENCES

[1] Wilma A Bainbridge, Justin W Hart, Elizabeth S Kim, and Brian Scassellati. 2011. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics* 3 (2011), 41–52.

[2] Marian R. Banks, Lisa M. Willoughby, and William A. Banks. 2008. Animal-Assisted Therapy and Loneliness in Nursing Homes: Use of Robotic versus Living Dogs. *Journal of the American Medical Directors Association* 9, 3 (2008), 173–177. https://doi.org/10.1016/j.jamda.2007.11.007

[3] Christoph Bartneck, Juliane Reichenbach, and Albert van Breemen. 2004. In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. *Proceedings of Design and Emotion* (2004).

[4] Cameron B. Browne, Edward Powley, Daniel Whitehouse, Simon M. Lucas, Peter I. Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. 2012. A Survey of Monte Carlo Tree Search Methods. *IEEE Transactions on Computational Intelligence and AI in Games* 4, 1 (2012), 1–43. https://doi.org/10.1109/TCIAIG.2012.2186810

[5] Jonathan Chan and Goldie Nejat. 2012. Social Intelligence for a Robot Engaging People in Cognitive Training Activities. *International Journal of Advanced Robotic Systems* 9, 4 (2012), 113. https://doi.org/10.5772/51171

[6] Veena Chattaraman, Wi-Suk Kwon, and Juan Gilbert. 2013. Interaction Style of Virtual Shopping Agents: Effects on Social Presence and Older Consumers' Experience in E-tail Sites. In *International Textile and Apparel Association Annual Conference Proceedings*, Vol. 70. Iowa State University Digital Press.

[7] Veena Chattaraman, Wi-Suk Kwon, Juan E. Gilbert, and Kassandra Ross. 2019. Should AI-Based, conversational digital assistants employ social- or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. *Computers in Human Behavior* 90 (2019), 315–330. https://doi.org/10.1016/j.chb.2018.08.048

[8] Luis Cobo Hurtado, Pablo Francisco Viñas, Eduardo Zalama, Jaime Gómez-García-Bermejo, José María Delgado, and Beatriz Vielba García. 2021. Development and usability validation of a social robot platform for physical and cognitive stimulation in elder care facilities. *Healthcare* 9, 8 (2021), 1067.

[9] Filipa Correia, Patrícia Alves-Oliveira, Nuno Maia, Tiago Ribeiro, Sofia Petisca, Francisco S. Melo, and Ana Paiva. 2016. Just follow the suit! Trust in human-robot interactions during card game playing. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 507–512. https://doi.org/10.1109/ROMAN.2016.7745165

[10] Filipa Correia, Patrícia Alves-Oliveira, Tiago Ribeiro, Francisco Melo, and Ana Paiva. 2021. A Social Robot as a Card Game Player. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 13, 1 (Jun. 2021), 23–29. https://doi.org/10.1609/aiide.v13i1.12936

[11] Filipa Correia, Tiago Ribeiro, Patrícia Alves-Oliveira, Nuno Maia, Francisco S. Melo, and Ana Paiva. 2016. Building a social robot as a game companion in a card game. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 563–563. https://doi.org/10.1109/HRI.2016.7451857

[12] Cristina A. Costescu, Bram Vanderborght, and Daniel O. David. 2014. The Effects of Robot-Enhanced Psychotherapy: A Meta-Analysis. *Review of General Psychology* 18, 2 (2014), 127–136. https://doi.org/10.1037/gpr0000007

[13] Ivan Culjak, David Abram, Tomislav Pribanic, Hrvoje Dzapo, and Mario Cifrek. 2012. A brief introduction to OpenCV. In *2012 Proceedings of the 35th International Convention MIPRO*. 1725–1730.

[14] Sam Cunningham-Nelson, Wageeh Boles, Luke Trouton, and Emily Margerison. 2019. A review of chatbots in education: practical steps forward. In *30th annual conference for the australasian association for engineering education (AAEE 2019): educators becoming agents of change: innovate, integrate, motivate*. Engineers Australia, 299–306.

[15] Roberta De Cicco, Susana Cristina Lima da Costa e Silva, and Francesca Romana Alparone. 2021. "It's on its way": Chatbots applied for online food delivery services, social or task-oriented interaction style? *Journal of Foodservice Business Research* 24, 2 (2021), 140–164.

[16] Juan Fasola and Maja J Matarić. 2013. A socially assistive robot exercise coach for the elderly. *J. Hum.-Robot Interact.* 2, 2 (June 2013), 3–32. https://doi.org/10.5898/JHRI.2.2.Fasola

[17] Martin Fischbach, Jean-Luc Lugrin, Michael Brandt, Marc Erich Latoschik, Chris Zimmerer, and Birgit Lugrin. 2018. Follow the white robot-a role-playing game with a robot game master. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. 1812–1814.

[18] Susan T. Fiske, Amy J. C. Cuddy, and Peter Glick. 2007. Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences* 11, 2 (Feb. 2007), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

[19] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 42, 3 (2003), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X Socially Interactive Robots.

[20] Elisabeth Ganal, Lenny Siol, and Birgit Lugrin. 2023. PePUT: A Unity Toolkit for the Social Robot Pepper. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1012–1019. https://doi.org/10.1109/RO-MAN57019.2023.10309447

[21] Elisabeth Ganal, Fiona Wiederer, Isabell Bachmann, and Birgit Lugrin. 2023. Who Would You Rather Ask for Help? A Comparison Between Two Chatbot Personalities in Information Management. In *International Conference on Human-Computer Interaction*. Springer, 334–353.

[22] Gene V Glass, Percy D Peckham, and James R Sanders. 1972. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research* 42, 3 (1972), 237–288.

[23] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. 2010. Assessing Acceptance of Assistive Social Agent Technology by Older Adults: the Almere Model. *International Journal of Social Robotics* 2, 4 (Dec. 2010), 361–375. https://doi.org/10.1007/s12369-010-0068-5

[24] Laura Hoffmann and Nicole C. Krämer. 2013. Investigating the effects of physical and virtual embodiment in task-oriented and conversational contexts. *International Journal of Human-Computer Studies* 71, 7 (2013), 763–774. https://doi.org/10.1016/j.ijhcs.2013.04.007

[25] Roland Imhoff, Jonas Woelki, Sebastian Hanke, and Ron Dotsch. 2013. Warmth and competence in your face! Visual encoding of stereotype content. *Frontiers in psychology* 4 (2013), 386.

[26] Peiyuan Jiang, Daji Ergu, Fangyao Liu, Ying Cai, and Bo Ma. 2022. A Review of Yolo algorithm developments. *Procedia computer science* 199 (2022), 1066–1073.

[27] Younbo Jung and Kwan Min Lee. 2004. Effects of physical embodiment on social presence of social robots. *Proceedings of PRESENCE* 2004 (2004), 80–87.

[28] Nina Jøranson, Ingeborg Pedersen, Anne Marie Mork Rokstad, and Camilla Ihlebæk. 2016. Change in quality of life in older people with dementia participating in Paro-activity: a cluster-randomized controlled trial. *Journal of Advanced Nursing* 72, 12 (2016), 3020–3033. https://doi.org/10.1111/jan.13076 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/jan.13076

[29] Martin Kaltenbrunner and Ross Bencina. 2007. reacTIVision: a computer-vision framework for table-based tangible interaction. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (Baton Rouge, Louisiana) *(TEI '07)*. Association for Computing Machinery, New York, NY, USA, 69–74. https://doi.org/10.1145/1226969.1226983

[30] Yukitaka Kawaguchi, Kazuyoshi Wada, Masako Okamoto, Takeo Tsujii, Takanori Shibata, and Kaoru Sakatani. 2012. Investigation of brain activity after interaction with seal robot measured by fNIRS. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. 571–576. https://doi.org/10.1109/ROMAN.2012.6343812

[31] Pouria Khosravi, Azadeh Rezvani, and Anna Wiewiora. 2016. The impact of technology on older adults' social isolation. *Computers in Human Behavior* 63 (2016), 594–603.

[32] Cory D. Kidd and Cynthia Breazeal. 2004. Effect of a robot on user perceptions. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, Vol. 4. 3559–3564 vol.4. https://doi.org/10.1109/IROS.2004.1389967

[33] Sara Kiesler, Aaron Powers, Susan R. Fussell, and Cristen Torrey. 2008. Anthropomorphic Interactions with a Robot and Robot-like Agent. *Social Cognition* 26, 2 (2008), 169–181. https://doi.org/10.1521/soco.2008.26.2.169 arXiv:https://doi.org/10.1521/soco.2008.26.2.169

[34] Min-Gyu Kim and Kenji Suzuki. 2014. Comparative Study of Human Behavior in Card Playing with a Humanoid Playmate. *International Journal of Social Robotics* 6, 1 (2014), 5–15. https://doi.org/10.1007/s12369-013-0184-0

[35] Adrian Krenzer, Michael Banck, Kevin Makowski, Amar Hekalo, Daniel Fitting, Joel Troya, Boban Sudarevic, Wolfgang G. Zoller, Alexander Hann, and Frank Puppe. 2023. A Real-Time Polyp-Detection System with Clinical Application in Colonoscopy Using Deep Convolutional Neural Networks. *Journal of Imaging* 9, 2 (2023). https://doi.org/10.3390/jimaging9020026

[36] Laura Kunold and Linda Onnasch. 2022. A Framework to Study and Design Communication with Social Robots. *Robotics* 11, 6 (2022). https://doi.org/10.3390/robotics11060129

[37] Kwan Min Lee, Younbo Jung, Jaywoo Kim, and Sang Ryong Kim. 2006. Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International journal of human-computer studies* 64, 10 (2006), 962–973.

[38] Iolanda Leite, Andre Pereira, Carlos Martinho, and Ana Paiva. 2008. Are emotional robots more fun to play with?. In *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*. 77–82. https://doi.org/10.1109/ROMAN.2008.4600646

[39] Darwin Li. 2024. The Cognitive Benefits of Playing Contract Bridge in Elderly Patients with Dementia. *International Journal of Psychiatry* 9, 2 (2024), 01–08.

[40] Jamy Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies* 77 (2015), 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

[41] Maria R. Lima, Meena Wairagkar, Niranjan Natarajan, Shilpa Vaitheswaran, and Rajkumar Vaidyanathan. 2021. Robotic Telemedicine for Mental Health: A Multimodal Approach to Improve Human-Robot Engagement. *Frontiers in Robotics and AI* 8 (2021), 618866. https://doi.org/10.3389/frobt.2021.618866 This

is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

[42] Wing-Yue Geoffrey Louie, Derek McColl, and Goldie Nejat. 2012. Playing a memory game with a socially assistive robot: A case study at a long-term care facility. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. 345–350. https://doi.org/10.1109/ROMAN.2012.6343777

[43] Birgit Lugrin. 2021. Introduction to socially interactive agents. In *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition*. Association for Computing Machinery, 1–20.

[44] Birgit Lugrin. 2022. Preface. In *The Handbook on Socially Interactive Agents: 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application*. Association for Computing Machinery, xxvi–xxxvi.

[45] Olimpia Pino, Giuseppe Palestra, Rosalinda Trevino, and Berardina De Carolis. 2020. The Humanoid Robot NAO as Trainer in a Memory Program for Elderly People with Mild Cognitive Impairment. *International Journal of Social Robotics* 12, 1 (2020), 21–33. https://doi.org/10.1007/s12369-019-00533-y

[46] Rui Prada and Diogo Rato. 2022. Socially interactive agents in games. In *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application*. Association for Computing Machinery, 493–526.

[47] Martin Prince, Renata Bryce, Emiliano Albanese, Anders Wimo, Wagner Ribeiro, and Cleusa P. Ferri. 2013. The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia* 9, 1 (2013), 63–75.e2. https://doi.org/10.1016/j.jalz.2012.11.007

[48] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[49] Byron Reeves and Clifford Nass. 1996. The media equation: How people treat computers, television, and new media like real people. *Cambridge, UK* 10, 10 (1996), 19–36.

[50] Robert Rosseter. 2014. Nursing shortage. *American Association of Colleges of Nursing* (2014), 1–6.

[51] Hiral Soni, Julia Ivanova, Hattie Wilczewski, Alexandra Bailey, Triton Ong, Alexa Narma, Brian E Bunnell, and Brandon M Welch. 2022. Virtual conversational agents versus online forms: Patient experience and preferences for health data collection. *Frontiers in Digital Health* 4 (2022), 954069.

[52] Pietro Tagariello, Paolo Girardi, and Mario Amore. 2009. Depression and apathy in dementia: same syndrome or different constructs? A critical review. *Archives of gerontology and geriatrics* 49, 2 (2009), 246–249.

[53] United Nations, Department of Economic and Social Affairs, Population Division. 2022. *World Population Prospects 2022: Summary of Results*. Technical Report UN DESA/POP/2021/TR/NO. 3. United Nations, New York. https://www.un.org/development/desa/pd/content/world-population-prospects-2022

[54] Willemijn M. van Dolen, Pratibha A. Dabholkar, and Ko de Ruyter. 2007. Satisfaction with Online Commercial Group Chat: The Influence of Perceived Technology Attributes, Chat Group Characteristics, and Advisor Communication Style. *Journal of Retailing* 83, 3 (2007), 339–358. https://doi.org/10.1016/j.jretai.2007.03.004

[55] Tibert Verhagen, Jaap Van Nes, Frans Feldberg, and Willemijn Van Dolen. 2014. Virtual customer service agents: Using social presence and personalization to shape online service encounters. *Journal of Computer-Mediated Communication* 19, 3 (2014), 529–545.

[56] Joshua Wainer, David J. Feil-Seifer, Dylan A. Shell, and Maja J. Mataric. 2007. Embodiment and Human-Robot Interaction: A Task-Based Perspective. In *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*. 872–877. https://doi.org/10.1109/ROMAN.2007.4415207

[57] Ao Wang, Hui Chen, Lihao Liu, Kai Chen, Zijia Lin, Jungong Han, and Guiguang Ding. 2024. Yolov10: Real-time end-to-end object detection. *arXiv preprint arXiv:2405.14458* (2024).

[58] Jiaji Zhu, Yushi Jiang, Xiaoxuan Wang, and Suying Huang. 2023. Social- or task-oriented: how does social crowding shape consumers' preferences for chatbot conversational styles? *Journal of Research in Interactive Marketing* 17, 5 (2023), 641–662. https://doi.org/10.1108/JRIM-01-2022-0007

[59] Aleksa Ćorović, Velibor Ilić, Siniša Ðurić, Mališa Marijan, and Bogdan Pavković. 2018. The Real-Time Detection of Traffic Participants Using YOLO Algorithm. In *2018 26th Telecommunications Forum (TELFOR)*. 1–4. https://doi.org/10.1109/TELFOR.2018.8611986