Behaviour & Information Technology

Adaptive Humanoid Robot Behaviour in a Serious Game Scenario through Reinforcement Learning

Submission ID	242625688
Article Type	Research Article
Keywords	Robot Behaviour Adaptation, Robot Personality, Reinforcement Learning, QL
Authors	ELEONORA ZEDDA, Marco Manca, Fabio Pater nò, Carmen Santoro

Adaptive Humanoid Robot Behaviour in a Serious Game Scenario through Reinforcement Learning

Eleonora Zedda[1*], Marco Manca[1], Fabio Paternò[1], Carmen Santoro[1]

[1]HIIS Laboratory, National Research Council (CNR) -Institute of Information Science and Technologies (ISTI) , Giuseppe Moruzzi, 1, Pisa, 56127, Italy.

*Corresponding author(s). E-mail(s): eleonora.zedda@isti.cnr.it; Contributing authors: marco.manca@isti.cnr.it; fabio.paterno@isti.cnr.it; carmen.santoro@isti.cnr.it;

Abstract

The study presents an adaptive technique that enables a humanoid robot to select appropriate actions to maintain the engagement level of users while they play a serious game for cognitive training. The goal is to design and develop an adaptation strategy for changing the robot's behaviour based on Reinforcement Learning (RL) to encourage the user to remain engaged. Initially, we trained the algorithm in a simulated environment before moving on to a real user experiment. Thus, we first design, develop, and validate the RL strategy in a simulated environment. Subsequently, we integrate the trained policy into the robotic system, allowing us to select the best robot actions based on the detected user state during real user test.

The RL algorithm was designed and implemented to determine an effective adaptation strategy for the robot's actions, encompassing verbal and nonverbal interactions. The proposed solution was first trained in a simulated environment and then tested with 28 users in a mixed-study design.

Keywords: Robot Behaviour Adaptation, Robot Personality, Reinforcement Learning, QL

1 Introduction

In the last decade, the interactions between people and robots have become more complex. Users start interacting and communicating with the robot rather than just operating it. This communication may use a number of modalities, including speech, touch, gestures, and vision, as the complexity of the systems rises. Modern interactive systems often use a combination of these modalities to communicate meaningfully. For example, a robot may coordinate its speech with its actions, considering visual feedback during its execution [1]. Additionally, with the expansion of social robots in society, these systems will impact users in several aspects of life, from providing assistance and cognitive training to taking part in collaboration tasks [2, 3]. Cognitive training tasks aiming to stimulate human resources, such as memory and attention, are among the most important challenges because they seek to stimulate sensitive users to improve their lives. From this perspective, social robots offer a potential solution to address the challenges associated with the monotony of cognitive training and enhance user engagement during repetitive tasks. Social robots, designed to assist human users through social and natural interaction, have gained increasing interest in recent years [4]. In order to facilitate natural interaction, researchers in social robotics have focused on robots that can adapt to diverse conditions and different user needs [5]. To enable machines to interact with users in a natural manner, the system must be able to identify or recognize the state in human behaviour and performance through the input modalities (e.g., speech recognition, user state recognition, gestures) and communicate through its output modalities (e.g. speakers, animations, and so on). Using machine learning (ML) techniques, the robot agent can sense its environment through the multi-modal input modules, plan the next action, and act using its output processing modules. Modelling and optimizing such interaction patterns necessitate learning and adaptation, which are essential features of any intelligent interactive system. ML methods are pivotal in solving this problem and optimizing interaction patterns. Among the various ML techniques, Reinforcement Learning (RL) is the most used framework for decision-making problems in which an agent interacts through trial-and-error with its environment to discover an optimal behaviour. Since interaction is a key component in RL and social robotics, it can be a well-suited approach for real-world interactions with physically embodied social robots [5]. Recent research has increasingly focused on robots that can adapt their behaviour to various human conditions to enhance user attention and engagement [6–10]. Adaptive robot interactions are important for providing comfortable and effective interactions with humans, fostering meaningful communication and building trust between users and robots [9, 11]. In the health domain, personalized and tailored robotic assistive systems can establish productive interactions with users, thereby improving the outcomes of therapy sessions [12–14]. Recently, interest has been rising in developing social robots that simulate social characteristics. This research dimension aims to develop natural and intuitive Human-Robot Interaction (HRI) to facilitate user acceptance, engagement and experience. One attempt is to design humanoid robots with social characteristics such as exhibiting a natural gaze, gestures, and personalities. Several studies have shown how personalities can influence different aspects of interaction, such as satisfaction, likeability, and other aspects that are central to making an effective

HRI. Indeed, researchers have claimed that personality is a key that triggers intuitive responses from users during HRI [15]. Moreover, personalities often shape the very nature of social relationships and influence the level of satisfaction derived from such interactions [16]. Earlier research has demonstrated that the personalities of social robots influenced user preferences [2] and also affected the perceived enjoyment of the interaction concerning the perceived intelligence and overall attractiveness of social robots [15]. Additionally, various social robot capabilities can help create an intimate and effective social interaction, such as expressing emotions, communicating through high-level dialogues, using natural cues, developing social skills and exhibiting distinctive personalities [4, 17]. Multiple studies [18–21] have indicated that social robots with personalities can enhance interaction, simulating the dynamics of human-human interaction during cognitive training administered by a human therapist.

Indeed, as the literature shows [6–10], adapting the social robot system with some social behaviours like robot personalities can result in a more productive, engaging HRI. Within this domain, some studies [9, 22] found that social robots' behaviour adaption to users' states improves a user's engagement and adherence to repetitive cognitive training. However, previous studies designed adaptive strategies focusing mainly on exploring robotic dialogue strategies [3, 11, 23]. In our study, we want to investigate suitable robot behaviour strategies composed of verbal and nonverbal parameters while exhibiting specific personalities on a social robot. To achieve this goal, we employ RL methodology to facilitate an adaptive approach for the robot in a serious cooking game for cognitive training of healthy and Mild Cognitive Impairment (MCI) subjects. First, we design, develop, and validate the RL strategy in a simulated environment. Then, we integrate the trained policy into the robot system, allowing us to select the best robot actions according to the detected user state in the real user test. The article is structured into the following sections: Introduction, where we describe the motivations and goals of this work; State of the Art, which reviews contributions in the field; Approach Section, which provides a detailed description of the key elements of our methodology; the User Modelling section, which describes how we classify and model the user state for both the simulation and real user test; the Robot Behaviour Modelling Simulation section that describes the key element of the QL algorithm used and the results of the simulation; the User Study section, where we outline the protocol and details of the conducted user test; the Results section contains the description of collected data, statistical analysis, and outcomes; and lastly we provide a Discussion, and a Conclusion sections.

2 State of the Art

In HRI, adaptation may handle different goals and address varying aspects of interaction. For example, learning adaptation with a focus on empathic, supportive strategies using a pet robot [24] or the effective behaviour of a tutoring robot [25], adapting personality in the assistive domain for post-stroke users [2], where the system also has to be able to deal with the user's disabilities.

Ritschel et al. [26] used a QL (QL) approach for adapting the Reeti robot to keep the user engaged during the interaction using user engagement. The key element in

Ritschel's work is user engagement. The article presents an approach based on RL, which gets its reward directly from social signals in real-time during the interaction to quickly learn about and dynamically address individual human preferences. In the specified scenario, the robot generates a description of a fantasy character chosen by the user and uses its limited knowledge base. This configuration restricts the interaction time to approximately three minutes, within which the robot can only provide brief descriptions of the fantasy character chosen. A limitation of this study is the choice of task. This limited interactive scenario limits the wide range of interactions and functionality that the robot could potentially provide. Another limitation is that the authors did not conduct a user test in this work.

Tapus et al. [12] designed an adaptive social robot system that customised a protocol through motivation, encouragement and companionship for users who have Alzheimer's disease. The robot dynamically adjusts game difficulty based on user performance, promoting users to higher levels if they surpass an Accepted Variation Band. The experimental group consisted of nine seniors over 70 with cognitive impairment or Alzheimer's disease from a senior facility. The social robot boosted cognitive attention and enhanced task performance in impaired patients through personalised encouragement and adaptive behaviour. However, the experimental group consisted of a limited sample of nine participants from the senior facility.

In another study by Magyar et al. [13], the authors employed QL to train the robot to learn a robotic conversation strategy to promote conversation with older adults considering the users' preferred topics and emotions. The authors developed a system considering the senior's motivation to speak and the older adult's emotions to implement a conversation strategy that can be adapted to the older adult's needs and preferences. The older adult's state was represented along two dimensions, each with three possible levels: emotion (negative, neutral, positive) and motivation (low, neutral,high). The robot's action space was represented by three actions: encouragement, question, and topic change.

The authors used RL to train the action selection model. However, the final action selection was made stochastically in the user test because the algorithms select short responses more often than the other two actions. Another limitation is the small number of test users (only three).

In Yuan et al. [3] work, the researchers develop an adaptive conversation strategy to answer people with dementia repetitive questions, follow up with new questions to distract PwDs from repetitive behaviour and stimulate their conversation and cognition using QL in a Markov Decision Process (MDP) problem. As a result of the simulation, they found that the implemented RL agent can learn the best policy within 30 interactions. The study has limitations, including the Markov decision process focusing only on the user's response relevance and overlooking latent variables like cognitive capability and user engagement level, which are important elements to consider in generating an effective adaptive strategy. Additionally, the manual coding of optimal policies in the demonstration, performed by a researcher simulating a user, may hinder long-term adaptive conversational robot development. For this reason, it is important to test the system with users. The the absence of a user test performed and the assumption of constant user commitment throughout HRI dialogue are also noted limitations.

Hemminghaus et al. [6] studied if a social robot can learn from a task-oriented interaction with a human user and how to employ different social behaviours to achieve interactional goals under specific situational circumstances. The system is applied in a cognitive training scenario in which the Furhat robot assists a human player in solving a memory game by guiding the user's attention to target objects. The behaviour generation process is made adaptive using QL. The analysis of the evaluation study results shows that the robot's behaviour in the learning condition achieved better task assistance than the behaviour in the random condition, even though the participants in the learning condition did not ask for help as much as in the random condition. This study presents some limitations, mainly because the system is not autonomous but implemented with the Wizard of OZ technique. The system does not detect the user state but is a researcher who decides when to intervene to perform an assistive action. Another limitation is that they tested the cognitive training system designed for older adults with students.

Previous studies [3, 6, 13, 26] designed adaptive strategies focusing mainly only one aspect of the game difficulty setting or focusing mainly on exploring limited robotic dialogue strategies rather than considering also non-verbal strategies. Social robots' natural conversational and gesture abilities should not be limited to short basic task-related sentences. However, the robots should be able to engage users in the interactions with entertainment, stimulating and encouraging the user in case it is in a "bad mood" (for example, by varying different robot behaviours that can have different effects on the user), and to re-engage the user during the interaction. Furthermore, some previous studies [6, 27] have implemented the system using a Wizard of OZ setup to decide when a robot should take a specific behaviour and to classify the user's state. On the contrary, the system proposed in this work allows the robot to decide completely autonomously what behaviour to perform depending on the user's state.

In this study, we aim to identify effective robotic dialogue and behavioural strategies that include both verbal and non-verbal parameters, manifesting distinct personalities, using QL (a model-free RL algorithm). In addition, the same vocal feedback and animation are not reproduced for the same behaviour, but the robot can choose from multiple types of feedback and animations specific to that behaviour. This was done to mitigate the repetitiveness of the robot's actions and obtain a more "natural" robot behaviour, also using an adaptive strategy to keep the user engaged and attentive during cognitive training.

To the best of our knowledge, the work presented is one of the first to use an adaptive robot behaviour strategy for social robots (such as the Pepper robot) exploiting verbal and nonverbal parameters in serious game scenarios. The strategy applied can support adaptive interactions with two types of users with various cognitive capabilities and engagement levels.

3 Approach

RL is a methodology already used to investigate the effective behaviour of social robots engaging with different typologies of users such as older adults [5, 28]. A crucial aspect is ensuring the robot consistently exhibits appropriate behaviour aligned with the user's state. Hence, the robot must be able to recognize the user's state and modify its responses and behaviour accordingly, accommodating individuals with varying cognitive skills.

We employed the Q-Learning algorithm (QL) [28] to facilitate an adaptive approach for the robot to be applied to a serious game and demonstrate its possibilities. As we will see in further detail later, the first simulated interaction phase using QL was used before deploying the trained learned policies onto the physical Pepper robot. The following subsections describe the key elements of the proposed approach.

3.1 Robot Personalities

Several studies [18–21] have focused on how to improve robot usability and user engagement in healthy and cognitive training scenarios by promoting a more natural HRI. In health and cognitive training scenarios, several contributions have suggested that humanoid social robots must have intimate and effective social interaction with humans [4, 29–31]. Various robot capabilities can be identified to create an intimate and effective social interaction, such as expressing emotions, communicating through high-level dialogues, using natural cues, developing social skills and exhibiting distinctive personalities [4, 17]. In the latter regard, various studies [2, 32, 33] have addressed the characteristics of personalities and have observed a close relationship between personalities and the modes of interaction between humans and robots [32, 34], suggesting that adding personality to a robot can make interactions more consistent and heighten user engagement and user experience [35]. Moreover, different studies [2, 15, 33] found that a robot with different personalities can simplify the interaction, which is particularly useful especially when e.g. users are older adults that carry out cognitive training exercises.

In particular, emerging humanoid robots may open up new possibilities in more effectively engaging MCI older adults during repetitive cognitive training [36]: beyond simplifying the interaction, including in the cognitive training, a robot exhibiting various personalities could enhance user engagement in therapy sessions, offer emotional feedback and foster a more natural and social behaviour surpassing their conventional perception as mere instrumental tools. However, to maximise the expressiveness of robots' behaviours and facilitate HRI in cognitive training settings, the interaction between the robot and the user should not rely only on verbal communication [37–39] but also considers nonverbal cues.

Several schools of thought regarding human personality exist. One of the most popular perspectives is trait-based. The traits are viewed as the primary mechanism by which personality manifests. A trait in personality can be defined as "a component or distinguishing characteristic of an individual's stable personality across time and external situations" [40]. Personality traits can predict an individual's attitudes and behaviour and have been identified as an essential facilitator of HRI.[41]. Although there are numerous traits models, the most used in the HRI community is the Big Five personality traits [41–43]. The Big Five personality traits include extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience [40, 44, 45]. The work focuses on the extraversion and introversion personality traits for the following reasons.

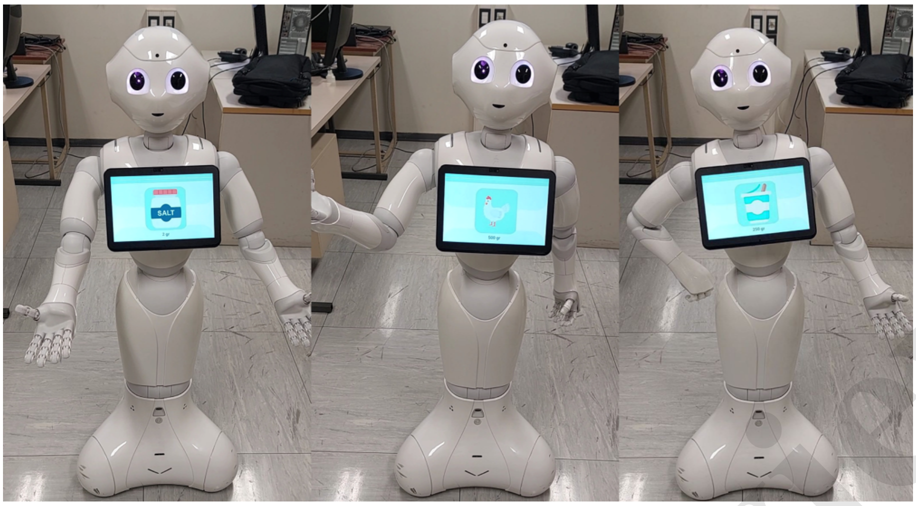
- a considerable amount of research indicates that extraversion is the most observable dimension among the Big Five Factors [16, 46–48];
- extraversion-introversion dimension is proven to be important in HRI [15, 27, 49], and affects users' quality of life and satisfaction during the interaction [50, 51]. In particular, the researchers K. Isbister and C. Nass [49] found that extraversion-introversion is the salient dimension in non-verbal cue research;
- verbal and non-verbal parameters used to represent extraversion and introversion are defined in the literature and can be efficiently modulated in robots [15, 27, 49, 52–56];
- the personality of a robot can provide users with better affordance, which makes it intuitive and natural for the users to understand the robot's behaviours [15, 57, 58].

To this aim, we identified a set of relevant parameters(seven verbal and three non-verbal cues). The verbal cues encompass pitch variation, volume, speech rate (measured in wpm = words per minute), pauses, dialogue style, syntax and sentences, while, the non-verbal ones involve the manipulation of animations, speed movements, and motor movements. Detailed information regarding these cues, refined through user tests conducted in different studies [59, 60] , is provided in Table 1. In summary, to simulate reserved behaviour, the introverted robot's vocal characteristics were adjusted by lowering its tone, slowing down its speech rate, slightly reducing its volume, utilising more frequent pauses, and using a kind and polite style of dialogue; the syntax comprises complex and indirect phrases. Conversely, the extraverted robot's vocal attributes were enhanced by raising its tone, increasing its speech speed, amplifying its volume, using shorter and less frequent pauses, with a more direct attitude and exclamations, the syntax used by the extravert contains more adjectives, prepositions, and articles, the phrases have a more varied vocabulary rich in adjectives and adverbs [32, 54, 61]. Regarding non-verbal parameters, in the extraversion condition, the robot's animations are more expansive, with broad animations to simulate openness toward the user. The animations generated in this condition usually involve elbows and hands moving away from the body using larger angles. For the introvert condition, the robot's animations tend to be more limited and contained. The animations generated in this condition usually involve the arms positioned close to the body, determining smaller angles in the introvert.

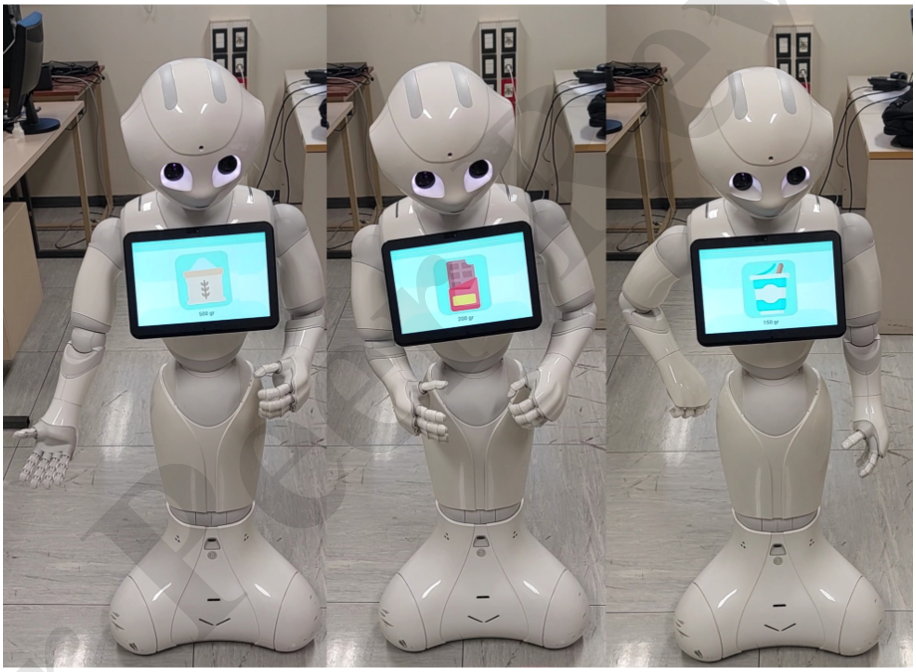
In the real user test, which will be described in Section 6, the caregiver has to select one of the two personalities integrated into the robot and remain constant during the entire session. The personality can be chosen according to the scope of the research study designed. According to the literature, to maximize the user's attention and engagement during the training session, we introduce the RL methodology to adapt the robot's behaviour according to the user state.

3.2 Reinforcement Learning

RL allows an agent to learn interactively how to behave in an unknown environment by observing the outcome of the performed action and the 'reward' it receives from



(a) An example of extravert animation performed during recipe explanation.



(b) An example of introvert animation performed during recipe explanation.

Fig. 1: An example of animations for extravert personality (Figure 1a), for introvert personality (Figure 1b).

Table 1: Final personality parameters identified for extravert and introvert personality.

Robot feature	Extravert	Introvert
Verbal Cues		
Pitch variation	~90% of maximum	~70% of maximum
Volume	~90% of maximum	~80% of maximum
Speech Rate	~180 wpm	~160 wpm
Pauses	400-600 ms	700-900 ms
Dialogue style	More direct, informal explicit, assertive	Polite, formal implicit reflexive
Syntax	present tenses more adjective and preposition direct construction	use past tense form, conditional sentences less adjective indirect construction
Sentences	direct appreciation emphatic, direct sentences	indirect appreciation less emphatic, hesitation sentences
Non-Verbal Cues		
Animation	animation with big angles more dynamic	animation with smaller angles less dynamic
Speed movements	faster movements more dynamic	slower and longer movements less dynamic
Motors movements	outward-directed movements, faster trajectory laterals, diagonal and forward	inward-directed, forward movements slower trajectory mainly backward

the environment. Thus, a robot applying RL can continuously improve its behaviour. However, there are some challenges when implementing RL in social HRI. One is the fact that learning the most appropriate robot behaviour by interacting with real users could require a non-negligible amount of interaction time, which can be tedious and impractical for users, resulting in users' fatigue and loss of interest [5]. Another disadvantage is the difficulty of finding a large sample of users needed to train the RL algorithm. A feasible alternative to cope with these challenges is the possibility of using a simulated world for the training phase before deploying the algorithm on the real robot. This has several advantages: first, it allows the agent to learn repeatedly, which would otherwise be very expensive; in addition, simulated environments can run much faster, thus permitting the learning agent to make more learning experiences at the same time; lastly, in the simulated environments, the number of users necessary to train the algorithms is not a barrier, and allows for avoiding time loss, expensive costs, and difficulties in recruiting a sample of sufficient size for the training. However, using simulated data can also present difficulties because it requires a policy showing success across different users' simulations and working in the real world. Indeed, simulating the real world can be difficult, especially when it comes to modelling relevant human behaviour (because simulating humans requires a predictive model of human interactive behaviour) and modelling the uncertainty of the real world.

To address these issues, we identified various parameters to better characterize user behaviour in a simulated environment using QL, a widely used approach to solve the Markov Decision Processes without estimating the environment model, and also using the result of preliminaries studies [3, 5, 62, 63]. Many researchers have applied it to the behaviour adaptation problem [3, 9, 13, 26, 64–68], namely the issue of selecting behaviours to maximize a person's engagement and performance. The advantage of using QL is that it is "model-free", meaning that it does not require a preexisting model: this is useful for behaviour adaptation because human reactions (i.e. which represent the 'rewards') are difficult to define as a model [26, 64]. Thus, we judged that QL was a suitable choice for a robot adaptation system applied to a serious game for cognitive training, as it can learn to make optimal decisions (the best robot action) based on experience gained through trial and error without having prior knowledge.

3.3 Serious Game

In our work, we chose a serious cooking game scenario to manifest the effectiveness of the robot behaviour adaptive system with robot personalities. Serious games have become increasingly prevalent in supporting and improving the assessment of functional and cognitive abilities and providing alternative solutions for patients' treatment, stimulation, and rehabilitation [69]. In particular, they can also positively affect cognitive abilities, as shown in prior studies [36, 70]. As such, we adopted a serious game as a pertinent scenario for cognitive training among the older adult population, recognizing it as a fitting domain for applying different robotic personalities to support older adults and enhance their level of engagement during therapy sessions.

The serious game focuses on stimulating various cognitive abilities, including attention, visual memory, and short-term memory. It takes the form of a cooking activity simulation, where participants undertake exercises to identify ingredient sequences, weights, and typologies in recipes.

The serious game unfolds through five distinct states: introduction, recipe instruction, question, answer, and ending feedback. At the beginning (introduction phase), the robot initiates the interaction by greeting the user and inquiring about readiness to play. In this state, the robot adopts a neutral personality, using the default parameters set in the robot. This neutral personality is held until the personality selection is shown at the end of this introductory phase. When the actual cooking game starts, the robot exhibits its personality, presenting and vocally synthesizing the ingredients specified in the chosen recipe ("recipe instructions" state). Notably, the robot underscores the importance of the sequential order and weight of ingredients during the recipe instruction phase.

After that state, the game progresses to quizzes ("question" state), requiring users to use their visual attention and working memory skills to identify and select the correct ingredients from a set of options. Then, users must choose the correct answer from four options ("answer" state). Users can interact with the system throughout the process through voice interaction. After the user answers the question, the robot provides reinforcement feedback to encourage the user to continue the game. Once users complete the game, they enter the "ending feedback" state, which displays the game's score (bronze, silver, gold). The robot's final feedback varies according to the number of correct answers provided by the user; also, it varies according to the robot's personality (by varying verbal, visual and non-verbal parameters).

4 User Modelling

The initial step in designing a QL system, is defining a classification of user states. Within our user modelling framework for QL, we differentiate between two user profiles: MCI and healthy individuals. This distinction allows for a more general classification of user states. The user state is composed of the user engagement and the user answer. Drawing from established theories on engagement [71], we classify user engagement based on gaze direction and smile. While the user answers, detect the correctness or wrongness of the user's answers.

This classification scheme remains consistent across real-user tests and simulations.

User Engagement

A person's engagement is defined as being occupied or involved with an external stimulus [71], i.e. the robot adaptation behaviours (i.e. robot actions) in our case. The engagement is conceptualized as having different dimensions [71]. In our work, we classified the level of engagement by considering the level of attention, measured as gaze direction, and the attitude toward the stimulus, measured as smile state. We determine the user engagement level by combining these values (user gaze and user smile).

Table 2 shows how the parameters of gaze direction and smile state are used to categorise the level of engagement into three levels.

Table 2: Value used to identify the user engagement level.

Characteristics	Feature used	Values	Meaning
User Attention	gaze direction	1	look at the robot face
		2	looking up
		3	looking down at the tablet
		4	looking left
		5	looking right
User smile	facial expression	1	not smiling
		2	smiling
		3	broadly smiling

The value "looking down at the tablet" in Table 2 is optional, depending on the structure of the robot feature (i.e. tablet). The values presented in the column "Values" in Table 2 are not ordered by importance but by meaning. Table 3 shows how we evaluate the engagement level by combining the user attention (as modelled using user gaze direction) and user smile state presented in Table 2. This categorization is used for the simulation and the user study. For example, the low engagement level is

Table 3: Values of user engagement classified by the user smile and user gaze direction.

Engagement Level	Attention Value	Smile Value
Low	2/4/5	1
Intermediate	1/3 or 2/4/5	1/2
High	1/3	2/3

characterized by a user not smiling and not looking at the robot. On the other hand, the high engagement level is characterized by a user smiling or broadly smiling and looking directly at the robot.

User Answer

In cognitive training, parameters such as the number of correct and incorrect answers play a crucial role in evaluating the cognitive and engagement state of users. The user's answers are marked as correct (marked as "1") or incorrect (marked as "0"). The engagement level and the user answer were combined to classify the user state during the simulation and the real interaction with the robot (described in detail in Table 4).

5 Robot Behaviour Modelling Simulation

Figure 2 describes the training used by the simulation process to get the most suitable "policy"(or behaviour) for a robot interacting with different user profiles (either MCI or healthy older adults) during the simulation. We incorporated the QL algorithm to keep the user engaged throughout the serious game by using the robot's adaptive behaviour. Indeed, typically, during cognitive training, the user is exposed to a series of repeated tasks that may create a high risk of dropping out of therapy and generate adverse conditions in users. To offset these problems, the robot should be able to detect the user's state and propose the most appropriate behaviour to engage users and stimulate their cognitive activities. If a user decreases the level of engagement, the robot should also be able to identify that state and adapt its action to re-engage him/her.

We will now describe in detail the main elements (in Figure 2) used in the simulation of the system, examining in particular how states, actions and rewards were defined. Since the robot should learn through its experience while interacting with the user, we endow it with an QL algorithm that can adapt without a previously specified model. In QL, the policy is formulated as a $Q(s, a)$ matrix, where s is the state of the environment (i.e. the user state defined in Section 4)) at a given time, and a is an action the robot uses to shape the environment. The state space consists of three dimensions: i) the engagement level, categorized as a value between

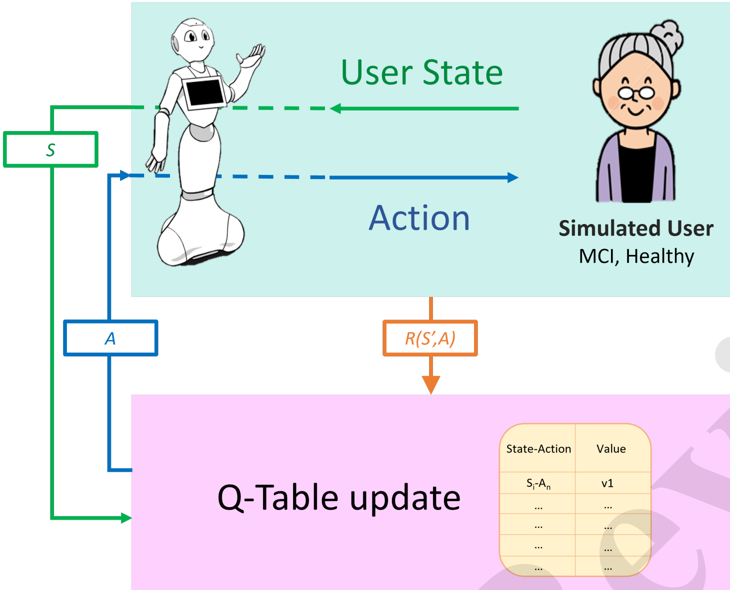


Fig. 2: Simulated interaction for training Q-Table. Where S is the current user state, A the action performed by the robot, S' is the user state after the robot action A, and the reward $R(S', A)$ are the weights given to the robot A depending on the user state S due to A itself. In the representation of the Q-table, there is an example of the mapping state-action (e.g. $S_i - A_n$) in which ($i \in [0-5]$) and ($n \in [0-2]$) with a value ($v1$) obtained with the update of the Q-table using the Bellman equation.

$E = \{Low(L), Intermediate(I), High(H)\}$; ii) the probability of answering correctly or wrongly according to the engagement level; iii) and the probability of changing the engagement level during the interaction as a consequence of a robot action, will be described in detail in Section 5.1.1. The action space consists of three possible actions represented as $a = A_0, A_1, A_2$ defined as the types of behaviours the robot provides. The robot agent learns its policy by interacting with the simulated users. At that stage, the Q-table is updated according to the standard equation [5]:

$$Q_{\text{new}}(s, a) = (1 - \alpha) \cdot Q(s, a) + \alpha \cdot [R(s, a) + \gamma \cdot \max Q(s', a)] \tag{1}$$

where s and s' are, respectively, the current user's state and the expected next user's state, a is the robot action, $R(s, a)$ is the reward function, $0 \leq \gamma \leq 1$ is the discount factor, and $0 \leq \alpha \leq 1$ is the learning rate. $Q(s, a)$, $Q(s', a)$, $Q_{\text{new}}(s, a)$ represent respectively the Q-value associated to the (s, a) pair, the estimate future value, and the update Q-value for the (s, a) pair. The expected next state s' corresponds to the

state for which the $Q(s', a)$ is maximum. Finally, the discount factor γ and the learning rate α determine, respectively, the importance of the future rewards in the learning process and how fast the model learns.
The probability of answering correctly according to the engagement level and the probability of changing the engagement level during the interaction are considered in $R(s', a)$, as explained later in Section 5.3. The standard equation converges to the Bellman Optimality Equation [72].

At each interaction step, the agent detects the user state s' and then performs an action a following the Q-table. Based on the outcome, the robot agent receives a reward $R(s', a)$ related to the new user's state, and the Q-values are updated. The simulation process goes on until the Q-values converge to the optimal values, and the learning process stops. During each interaction with the user, the agent needs to learn the most suitable policy and the mapping from states to actions that maximize the values contained in the Q-table for each state-action pair. These actions have to be selected among those the robot can carry out to stimulate the different user profiles while maintaining a high level of engagement. The details of the simulation process and the experiment are given in the following sections.

5.1 State

A state s is defined according to a user's state during a serious game scenario in a simulation environment. The state represents how we model the user in our simulation; in particular, the user state is composed of a combination of three user engagement levels and the user answer. The engagement level of each user profile (MCI and healthy) is categorized as low (L), intermediate (I) and high (H). To evaluate the user state, the user answer and the engagement level are combined to identify six user states. Table 4 shows how the identified user states are categorized according to the combination of the user engagement and user answer).

Table 4: Set of user states used for classifying the user state for simulated and real tests.

State ID	Engagement	Answer
0	L	0 (wrong)
1	L	1 (right)
2	I	0
3	I	1
4	H	0
5	H	1

To simulate user behaviour within the QL framework in the simulated environment, we introduce dynamic elements to capture the probabilistic nature of user response and engagement level, which is described in the following section 5.1.1.

5.1.1 Simulate MCI and Healthy users.

We model the user's behaviour of two kinds of users (an MCI user and a healthy user) to evaluate our robot behaviour modelling that captures different user abilities. These user models describe the user skills by using some parameters:

- the user's cognitive impairment (MCI or healthy);
- the user's engagement level;
- probability giving the correct and wrong answer;
- probability of changing the engagement level.

The individual's engagement with and without MCI can be influenced by robot attributes and the individual's attributes [73]. We assume the individual's cognitive capability and engagement level influence the probability that the user would provide a wrong or correct answer, as modelled in Table 5.The values chosen have been defined with the help of different psychotherapists and also based on previous studies in working with MCI and healthy older adults [59, 60].

Table 5: Modelling the probability for a user to give a wrong or a correct answer according to the type of user (MCI vs. healthy) and the user's engagement.

User Simulated	Engagement Level	% Answer Right	% Answer Wrong
MCI user	Low	30%	70%
	Intermediate	60%	40%
	High	80%	20%
Healthy	Low	60%	40%
	Intermediate	80%	20%
	High	90%	10%

As described in Table 5, a person with a lower engagement level will show a lower correctness answer probability; likewise, we expect a person with lower cognitive capability will have a higher possibility of answering wrong with a low engagement level compared to a healthy user. We also assume that the user's engagement state will change depending on the robot's actions. For this reason, we consider the probability of a changing in user engagement according to values shown in Appendix Table A1(for MCI users) and Appendix Table A2 (healthy users). The values in each cell denote the probabilities of modifying the engagement level when the corresponding action is taken in the given state. For example, we observed that for an MCI user in a state with a High engagement (Table A1, From High (H) to →), if an engaged action (Action0) is taken, then in 2 out of 10 cases the user engagement level may decrease to Low (non-attentive user). The rate mentioned refers to the probabilities of modulating the user engagement level in the MCI simulation.

The values in Tables A1, A2 were identified to simulate a real user interacting with a robot presenting a serious game and were defined with the help of a psychotherapist and also according to our experience with MCI and healthy older adults.

5.2 Action

The Actions correspond to the robot's behaviour (expressed in terms of verbal feedback, vocal parameters, animation and motor movements) following the for the robot personalities. In other words, the actions produce different interaction modality configurations. During the serious game application, the robot may take three typologies of actions.

In particular, $\mathbf{A}_0$ generates an engaging and energetic behaviour: for example, in the extravert personality, the robot will provide enthusiastic feedback such as: "My gosh! That is the correct one! You are trying hard!" slightly increasing the speech speed, volume, and pitch, also having a quite more dynamic and extensive animation and with more significant motor movements.

$\mathbf{A}_1$ generates a neutral robot behaviour: an example is: "Is the right answer!" with neutral vocal parameters defined by its personality and with a basic animation according to the robot's personality performed.

$\mathbf{A}_2$ generates a stimulating behaviour: an example is, "Right answer! Let us continue with this focus!" and more compact animations. All the parameters (verbal and non-verbal) used to define the three types of actions are based on the parameters of the two personalities described in [59, 60].

The actions generated by both personalities have the same purpose: to be either more energetic, stimulating, or neutral. However, each personality will manifest the three actions by slightly modifying the parameters defined in Table 1.

5.3 Reward

The reward function $R(s, a)$ specifies which states are more or less desirable, and the agent's goal is to choose actions to visit states with large positive rewards and avoid states with negative rewards. We have developed an immediate reward to stimulate the user to maintain a high level of engagement and concentrate on the questions. The weights specified in the algorithms are chosen to find the adaptive robot behaviour that is the most suitable to maintain a high user engagement level. The reward $R(s, a)$ depends on the parameters specified as it follows:

$$R(s,a) = \begin{cases} Engagement: & H = +10; \\ & I = -2.5 \\ & L = -5 \\ & \\ Answer: & Correct = +10 \\ & Wrong = -5 \end{cases} \tag{2}$$

In particular, a positive reward of +10 is given to the robot when the user exhibits a high level of engagement. Conversely, if the user's engagement level is intermediate, a slight penalty of -2.5 is applied to the reward. Furthermore, if the user has a low engagement level, a negative reward of -5 is assigned. This design aims to prevent the user from entering a state of low engagement. Another parameter considered in formulating the reward function is the user's responses. Specifically, if the user answers a question correctly, the robot action receives a reward of +10; otherwise, a penalty of -5 is given. These reward values are then weighted by multiplying them with the probabilities derived from the user profile. This approach encompasses the robot's actions and the user's cognitive state during the simulated interaction. Furthermore, to expedite the algorithm's convergence, a slight penalty of $-0.05t$ is imposed after each iteration t performed during the session. This is done because we would prefer our agent to reach a high user engagement level faster and maintain it until the end of the session, which is especially important during repetitive cognitive training.

5.4 Learning Experiments

In this section, we present the simulated experiment we did to obtain the optimal Q-values, which are then provided to the robot to choose which action a is the most suitable for that user state s, and the results of the user's interaction with the robot behaviour adaptation. Based on the definition above for QL's three key elements (state, action, and reward function), we have trained our QL model in Python by simulating 100 users, and a session of 35 iterations for each user. The discount factor and learning rate are set to 0.2 and 0.8, respectively, to ensure a quick and stable convergence of the training.

We initialize at zero a Q-table matrix of dimension $S \times A$, where $S = 6$ is the state space, and $A = 3$ is the action space. The simulation loops over the 100 users, and for each user, it loops over 35 possible interactions with the robot agent. At the first iteration, the user state s is generated randomly, and it corresponds to the user state identified by the robot agent at the beginning of the session with the user. The action a performed by the agent is chosen by applying the ϵ-greedy policy, which is a method to balance exploration (i.e. choose randomly the action a) and exploitation (i.e. choose the action a exploiting the knowledge of the Q-table $Q(s, a)$) by randomly choosing between them. The ϵ refers to the probability of choosing to explore, while 1-ϵ is the exploitation, the probability of choosing the action as $argmax_s Q(s, a)$, i.e. the action correspondent to the maximum Q-value given the state s. During the first iterations, we do not have enough knowledge about $Q(s, a)$, so the ϵ-greedy policy starts with $\epsilon = 1$, i.e. full exploration, and decreases exponentially to $\epsilon = 0.2$ as the number of iterations increases in order to exploit the gained knowledge. Given $s \in S = \{s_0, s_1, s_2, s_3, ..., s_5\}$, which correspond to a given engagement level E and answer state, as introduced in Table 4, and $a \in A = \{\mathrm{A}_0, \mathrm{A}_1, \mathrm{A}_2\}$, we assign the expected reward $R(s, a)$ and update the value of $Q(s, a)$ according to Equation 1. The expected next state s' is obtained as the $argmax_a Q(s, a)$. Then we assume s' as the state s in the next iteration. After a session of 35 iterations, we get the update $Q(s, a)$ that we use as input for the next simulated user. The pseudo-code shown in the following Algorithm 1 shows the logic of the QL training based on simulation.

The average reward, i.e. the average Q-table over the 100 simulated users per number of interactions, is used to assess our model's performance and evaluate the convergence of the training. From Figure 3, we observe that the algorithms for each user profile converge to the most suitable "policy" after 15 iterations. Thus, reaching the convergence, additional training will not improve the model. In comparison, the random policy, which chooses the actions and states randomly, does not reach convergence within 35 iterations, and on average, the performances are worse.

State $s \in S = \{s_0, s_1, s_2, s_3, ..., s_5\}$;
Action $a \in A = \{\mathrm{A}_0, \mathrm{A}_1, \mathrm{A}_2\}$;
Initialize $Q(s, a) = 0$;
for user < 100 **do**
    for $it < 35$ **do**
        if *it == 0 and user==0* **then**
            $s \leftarrow$ random;
        end
        $a \leftarrow \epsilon$−greedy$_s$;
        $r \leftarrow R(s, a)$;
        $(s') \leftarrow \arg\max_a Q(s, a)$;
        $Q_{\text{new}}(s, a) \leftarrow (1 - \alpha) \cdot Q(s, a) + \alpha \cdot [R(s, a) + \gamma \cdot \max Q(s', a)]$;
        $s \leftarrow s'$;
        $Q(s, a) \leftarrow Q_{\text{new}}(s, a)$;
        it++;
    end
    user++;
end

Algorithm 1: Pseudo-code of the QL training based on simulation.

The simulation for both user models reaches the convergence of performance around the same number of iterations. As expected, the average reward for the MCI user is lower than the average reward for the healthy user because the MCI user has more difficulty in keeping a high level of engagement during the session and in giving the correct answers, as observed during our preliminary studies on which we based our simulation of the MCI user through the engagement probabilities and the probability to give the correct answer. To prove the ability of the QL agent to adapt to eventual users' changes, we defined two additional extreme user profiles: an "Optimal" user scenario, represented by the green line in Figure 3. This user consistently provides correct answers following the robot's action, resulting in high engagement. Conversely, we simulate a user who consistently fails to provide correct answers, called "Inactive", leading to low engagement levels after each robot action. This user is represented by the red line in Figure 3.

The final Q-table provided by the optimal QL agent in each condition (MCI and healthy) is shown in Table 6, which provides the optimal action found for each of the six user states. From that Table, it is possible to see that MCI users need more stimulation than healthy ones. Stimulating or engaging actions are preferred to maintain a high or intermediate level of engagement for MCI users, whereas engaging actions should be preferred for healthy users. The trained Q-table obtained by the simulation is used in the Pepper robot application in the real user test explained in Section 6.

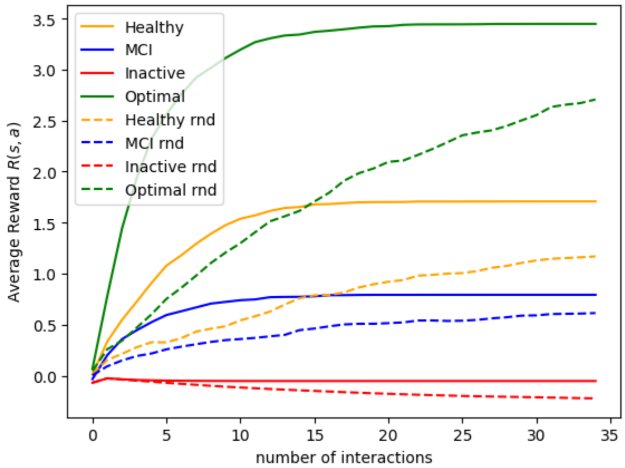


Fig. 3: The learning results of average return for User MCI in blue, Healthy in orange, Green for the Perfect and Red for the Inactive user. The solid and dashed curves represent the Q-learning and random selection learning results, respectively.

Table 6: Optimal Q-table results for the four user profile: M (MCI User), H (Healthy).

Action	A_0	A_1	A_2
State0			M H
State1		M H	
State2	H		M
State3	M H		
State4	M H		
State5	H		M

6 User Study

Before conducting the user test involving 28 participants, a preliminary study was undertaken with a psychotherapist specialized in cognitive training for older individuals with MCI. Specifically, the purpose of this meeting was to design the robot's adaptive actions, based on personality parameters, to maximize their positive effects while minimizing any potential negative impacts on users.

6.1 Pre-Study

This pre-study aimed to ensure that the design of the three actions, in particular animations and communication styles, would effectively stimulate users' attention and engagement without generating any negative feelings or judgments. Remote meetings were held in October 2022 with a psychotherapist of the local "Train the Brain" cognitive training clinic, who has expertise in cognitive training for older individuals with MCI and psychotherapy. Thus, we involved psychotherapists in the design process to identify which actions can positively impact users' attention and engagement, thereby being engaging, supportive, and motivating for users. The psychotherapist's expertise helps ensure that the selected actions align with psychological principles known to stimulate attention and engagement in individuals with MCI. Thus, three robot actions were designed following the two defined personalities. In other words, the three actions can be considered sub-behaviours of each robot's personality. With the help of the psychotherapist, we identified three verbal and three non-verbal cues as the most relevant features for the design of the three behaviours. All these parameters will follow the guideline for each robot personality identified in Table 1. For verbal cues, we modulate pitch, speech rate, and feedback goals, while for non-verbal cues, we manipulate gesture, gesture speed and motor movements. A summary of the cues tuned from the user tests performed can be found in Table 7.

6.1.1 Action Typology

The design of the robot's three adaptive actions is based on the two robot personalities described in section 3.1. Each of the two personalities possesses the three adaptive actions. This was done because previous studies have shown the positive effect of personalities on healthy and MCI users by increasing their UX and engagement[2, 15, 33, 36, 59, 60]. Adding these three adaptive actions following the defined personality parameters could help make the interaction with the robot more stimulating and interesting and also create a more natural interaction.

Engaging Action (A_0)

A_0 generates a more engaging and dynamic behaviour. To support this, some adjustments were made to various aspects of the robot's performance. First, the pitch and speech rate of the robot's voice increased by $\sim 5\%$. We include pauses to avoid that the robot's feedback is too fast and difficult to understand. Second, the animations employed by the robot were designed to be more rapid and dynamic for both personalities. Furthermore, the robot's movements were accentuated and made more visible during these actions. The feedback for this action is enthusiastic and motivating, aiming to acknowledge the user's effort and reinforce the correctness of their response. It may include phrases that express surprise, recognition of the user's hard work, and encouragement to continue performing well. Notably, all feedback generated aimed to avoid eliciting negativity or a sense of being judged in the user. Instead, the feedback was tailored to inform users about their responses while maintaining a positive intention. This principle was upheld consistently across all forms of feedback the robot offers, regardless of the action type or the personality involved.

Neutral Action (A_1) .

The action neutral (A_1) adheres to the predetermined behavioural design for both personalities without manipulating parameters for this specific type of action. The feedback for this action is straightforward and affirming, simply acknowledging that

Table 7: Parameters manipulation for design adaptive actions according to each robot personality

	Extravert		
	A_0 Engaged	*A_1 Neutral*	*A_2 Stimulating*
Pitch variation	95% of maximum	90% of maximum	80% of maximum
Speech Rate	$\sim$ 190 wpm	$\sim$ 180 wpm	$\sim$ 175 wpm
Feedback	Engaging	More direct	Stimulating
Gesture	gesture with bigger angles more dynamic and broadly	gesture defined by robot personality	slightly more closed and stationary animations
Speed	faster movements movements	neutral movements	slightly slowly movements
Motors	$\sim$ 20 cm	$\sim$ 15 cm	$\sim$ 10 cm

	Introvert		
	A_0 Engaged	*A_1 Neutral*	*A_2 Stimulating*
Pitch variation	75% of maximum	70% of maximum	55% of maximum
Speech Rate	$\sim$ 170 wpm	$\sim$ 160 wpm	$\sim$ 155 wpm
Feedback	Polite engaging	Polite	Polite Stimulating
Gesture	gesture with bigger angles more dynamic broadly	gesture defined by robot personality	slightly closed and stationary animations
Speed	slightly faster movements	neutral movements	slightly slowly movements
Motors	$\sim$ 15 cm	$\sim$ 10 cm	$\sim$ 5 cm

the answer given is correct without excessive praise or elaboration. It serves to validate the user's response and provide reassurance.

Stimulating Action (A_2)

The stimulating action (A_2) evokes a heightened sense of motivation and encouragement. Thus, adjustments were made to various aspects of the robot's performance to enhance the encouraging nature of the actions. Firstly, the robot's voice pitch was decreased by $\sim$ 5% from the neutral condition, with a reduction in speech rate. This decrease was decided to accentuate the feedback of encouragement and stimulation. Secondly, the animations utilized by the robot were purposefully designed to exhibit a slightly slower tempo and increased rigidity while maintaining a dynamic quality and avoiding excessive stillness. Additionally, the robot's movements were slightly

Table 8: Examples of verbal feedback used in the three typologies of actions for the correct and wrong answers state.

Action	Extravert Feedback	Introvert Feedback	State	Effects
A_0 Engaging	"My gosh! I agree with that! You are trying hard!"	"Ehm should be correct. It seems it is going well, ehm, come on!"	Answer Right	Encouraging Appreciation
A_1 Neutral	"Good! That's the right answer!"	"That should, ehm, be correct.."	Answer Right	Informational Neutral
A_2 Stimulating	"Good job! Let's keep that focus!"	"Um, that seems correct.. I think, um, maybe we can do better and better"	Answer Right	Stimulation
A_0 Engaging	"Is wrong, let's continue! Come on! We can do it!"	"Wrong, ehm..but I think we could give it, another try	Answer Wrong	Encouraging Appreciation
A_1 Neutral	"Oh, I'm sorry, the answer is wrong!!"	"Ehm, the answer seems wrong"	Answer Wrong	Informational Neutral
A_2 Stimulating	"Wrong! Come on, let's stay focused! Let's keep it going!"	"Is wrong, ehm, maybe we need to concentrate a little more! ehm, come on"	Answer Wrong	Stimulation

diminished in size and made somewhat less prominent, yet remained observable during these actions. The feedback for this action encourages the user to maintain focus and concentration. It may include phrases highlighting the user's effort, praising their performance, and motivating them to continue performing well. The psychotherapist's valuable input aided in devising the appropriate feedback mechanism for the robot across various action types. During the meeting, the psychotherapist suggested adjusting the feedback provided by the two robot personalities and highlighted the importance of providing encouraging feedback for both positive and negative answers. Table 8 shows examples of robot feedback with different personalities and action typologies.

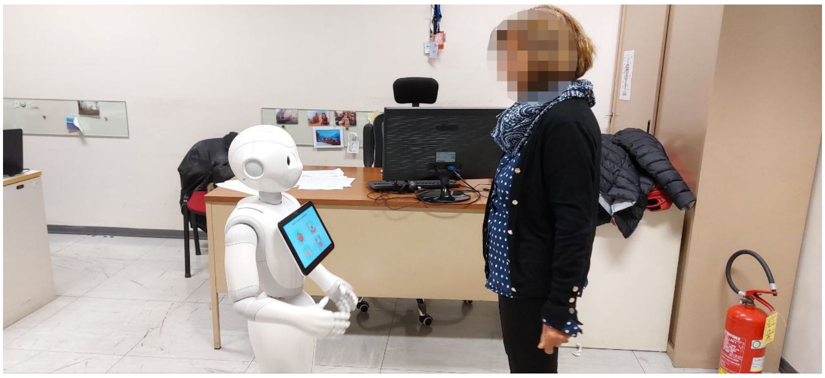


Fig. 4: An older adult during an adaptive session with Pepper exploiting extravert personality.

6.2 User Study

The study aimed to evaluate the effects of the QL-trained robotic strategy composed of verbal and non-verbal parameters while exhibiting specific personalities with adults and seniors. In this study, we address the following research questions:

1. **RQ1:** Can users discern between adaptive and random robot behaviour?
2. **RQ2:** Does the adaptive robot behaviour have an influence on user engagement and user experience?
3. **RQ3:** Does the adaptive robot behaviour affect user perception in terms of likeability, perceived intelligence, and anthropomorphism?

6.2.1 User Test Protocol

The test was conducted in a laboratory in December 2022. Participants were recruited through word of mouth and with notices posted inside the CNR and the University of Pisa. The enrolment requirements were to be at least 18 years old and Italian-speaking. For the test, the users interacted individually in the laboratory, standing in front of the robot at a distance of about 80 cm (see Figure 4). The experiment took an average of 45' for each user (test + questionnaires). A moderator was present and took notes of user feedback, user behaviour and any significant event occurring during the test. After the end of the test, the users were rewarded with some chocolates.

Participants. 28 (12 females) users between 20 and 72 years old (M = 43.1, SD= 16.77) were enrolled. Twenty-four had a university degree, and seven had a high school diploma. The majority, 78% (22 users), had no experience with robots. Only five had previously seen a Pepper robot on TV or the internet. Two had a Roomba robot at home, and one user knew the Asimo robot.

Test organization. The study employed a mixed design (between-subjects factor: robot personality) where all participants interacted with the two conditions(within-subjects factor):

- Random: the robot randomly chooses one of the three actions to perform with no adaptation condition.
- Adaptation: the robot selects the actions according to the user state detected and follows the Q-table trained.

To avoid the learning effect, all the conditions were counterbalanced, ensuring each participant experienced both conditions in a randomized order. half of the users (randomly selected) first interacted with the random robot, then with the adaptive one; the others did the opposite. Additionally, the robot personality was counterbalanced to obtain homogeneity in the sample for each condition. The robot's personality was counterbalanced to achieve homogeneity in the sample for each condition. In other words, each user interacted with only one personality under two conditions (random vs adaptation). Moreover, the recipes exploited during the two sessions were different to avoid the learning effect.

- *Step 1: Introduction.* At the beginning, participants were provided with an introduction to the main goals of the study, which included a presentation of the Pepper robot and an explanation of the tasks to carry out: to prevent any potential bias in the participants, the two robot conditions were not discussed throughout the duration of the session. Then, users signed a written informed consent describing the purpose of the research, the procedure of the study, duration, personal data processing information following the European Data Protection Regulation, the possibility to request the release of the data and how they are processed.
- Step 2: First session with a robot behaviour and questionnaire. In this step, there was an interaction with the robot showing the random or adaptive behaviour during a cooking game about preparing a recipe with 16 questions about the ingredients. Afterwards, the users had to compile a questionnaire about socio-demographic data, information about previous experience or familiarity with robots, six statements regarding the user robot adaptation perception (S.5, S.6), likeability (S.1, S.2, S.3) and satisfaction (S.4), the User Engagement Scale in short form (UES-short) and The Godspeed Questionnaire Series and some open questions [74] (see Appendix B). Socio-demographic data, including age, gender, education level, programming skills and prior experience with robots, were gathered to obtain a comprehensive profile of the user sample.
- *Step 3: Second session with a robot behaviour and questionnaire.* During the 2nd session, the application proposed the same game but involving a different recipe, while the robot exhibited the opposite condition. After the 2nd session, the same questions and questionnaires used in the 1st interaction were administered (excluding, those related to demographic information), also adding a question asking the preferred robot behaviour.
- *Step 4: Semi-structured Interview and Final Feedback.* In the semi-structured interview (Appendix B), we asked some questions regarding, i.e. whether users had perceived differences between the two types of robot's behaviour, whether users perceived differences in the way the robot speaks or moves, the likeability of the two types of robot's behaviour, and to what extent, in the user's perspective, the behaviour shown by the robot seemed to adapt or change throughout the interaction

Table 9: Research questions addressed for the adaptation study, with the data used to answer the research questions and statistical methods.

Research Questions	Data	Methods
RQ1	Adaptation Statements:	Paired t-test
Can users discern between adaptive and random robot behaviour?	S.5, S.6	
	Open Questions	Qualitative Analysis
	Semi-structure interview	Thematic Analysis
RQ2	UES-short	UES score calculation Paired Wilcoxon test
Does the adaptive robot behaviour have an influence on UE and UX?	Likeability Statements: S.1, S.2, S.3	Paired Wilcoxon test Paired t-test
	Satisfaction Statement: S.4	Paired Wilcoxon test
	Preferences Question Open Questions	Descriptive Statistic Qualitative Analysis
RQ3		
Does the adaptive robot behaviour affect user perceptions?	Godspeed questionnaire	Paired t-test

7 Results

In order to describe the research design and methodology used in the study, Table 9 provides a comprehensive overview of the research questions, the corresponding data sources for addressing them, and the methods employed to analyse the collected data.

7.1 RQ1 Can users discern between adaptive and random robot behaviour?

Table 10 presents the outcomes regarding the perception of the two robot conditions (see the questionnaire for the evaluation of robot adaptation in Appendix B3). The evaluation primarily focuses on whether users noticed any adaptive behaviour in the robot's feedback and movements. A paired samples t-test was conducted to investigate the influence of the robot's adaptive behaviour on this perception. To assess the statement, "I noticed some changes in the animations or movements performed by this robot", the normality of the data was examined using the Shapiro-Wilk's test. The test resulted in $W = 0.92, p = 0.06$, suggesting that a normality assumption could be made. Subsequently, a paired t-test was performed ($t = -3.62, p = 0.00011$), revealing a significant difference between the two conditions. Specifically, the adaptive condition (Mean=4.10, SD=1.08) was perceived as more adaptive regarding animations than

Table 10: Analysis of six statements for manipulation check for user perception of an Adaptive (A) and Random robot (N)

Category	M(SD) Adaptive Robot	Random Robot	Significant
S1.Likeability Feedback	**4.21**(0.67)	3.39 (0.97)	* $p < 0.005$
S2.Likeability Animation	**4.14**(0.74)	3.60 (1.08)	* $p < 0.005$
S3.Likeability overall HRI	**4.57**(0.56)	3.96 (0.85)	* $p < 0.005$
S4.Satisfaction	**4.21**(0.97)	3.86 (1.01)	ns
S5.Adaptation Feedback	**4.0**(0.97)	2.89 (1.23)	* $p < 0.005$
S6.Adaption Animation	**4.10** (1.08)	2.92 (1.55)	* $p < 0.005$

the random one (Mean=2.92, SD=1.55). For the statement "I noticed a change or increased stimulation in the way the robot gave me vocal feedback", the data were found to be normally distributed ($W = 0.94, p = 0.14$). Consequently, a paired t-test was conducted ($t = -3.07, p = 0.00047$), revealing a significant difference between the two conditions. In this case, the adaptive condition (Mean=4.0, SD=0.97) was perceived as more adaptive and stimulating through different verbal feedback compared to the random robot (Mean=2.89, SD=1.23).

7.1.1 Open Questions about robot adaptations

We analysed the answers that users gave to the open questions (see Appendix B Q.1-Q.6), to evaluate if users perceived the adaptation of the verbal and non-verbal behaviour in the two conditions (namely: the robot exhibiting adaptive behaviour versus the one showing a random behaviour). Users shared their insights on various aspects of the robots' speech and behaviour, as reported below.

Verbal behaviour. Several users reported noticing changes in the tone and speech rate of the adaptive robot behaviour during the test. In particular, they observed alterations in the adaptive robot's speech patterns when they made consecutive errors or provided multiple correct answers in a row. Some users indicated that they did not observe any noticeable change in the random robot. According to their responses, this group of users found the robot's speech and behaviour consistently changing throughout the interaction without any discernible patterns. Regarding the speech patterns, some users identified a slower rate and immediately after a higher speech rate in the random robot's responses. This generates incoherence in how the random robot behaves. These observations highlighted an incoherent communication style employed by the random behaviour robot, which deviated from the more dynamic and adaptive speech patterns observed in the adaptive robot. Some users noted that the random robot behaviour sometimes changes the tone of voice based on user responses. However, some users have found these responses incoherent.

Non-verbal behaviour. There were notable changes in the adaptive robot's behaviour for users. Users observed a shift in the robot's attitude when the interaction progressed: it became more amiable, with increased movement speed and intermittent bursts of animation. This transformation suggested heightened engagement from the robot, aiming to create a more interactive and responsive experience. Moreover, users noted that the adaptive robot used more conciliatory responses and animations, aligned and coherent with its verbal feedback. When users provided incorrect answers, the robot responded with empathy, using phrases and gestures conveying understanding, encouragement, and stimulation. These observations indicated that the adaptive robot aimed to establish a positive and supportive environment, leveraging verbal and non-verbal cues to create a more engaging interaction. The analysis of user comments reveals significant observations regarding the robots' animations and gestures. While interacting with the adaptive robot, several users observed changes in gestures and body movements. They noticed that the robot went from minor gestures to surprising actions when they responded well in a row but were still consistent during the interaction. These changes have contributed to a more immersive interaction with the robot. Users also noticed that the adaptive robot displayed specific gestures and movements in response to user actions. For example, when users gave incorrect answers, the robot moved backwards or sideways, exhibiting more "agitated" behaviour.

Conversely, when users provided correct answers, the robot moved more enthusiastically as it approached. This synchronisation between the robot's gestures and user responses created a dynamic and engaging interaction. Most users have observed that the robot animations have become more expansive and expressive, particularly during celebrations or encouragement. Furthermore, they noticed variations in the trajectories of the robot's movements, which became more spread and dynamic when the situation required it. Conversely, the majority of users interacting with the random robot did not notice any changes in animations or gestures. They perceived the robot's behaviour as consistently haphazard, with no significant variation during the interaction. Additionally, some noted that sometimes the animations did not align with what was being said or that the feedback did not seem to feature a consistent "human-like" behaviour.

Overall, users perceived greater stimulation and changes in the behaviour and speech of the adaptive robot compared to the random robot. The adaptive robot demonstrated adaptability by adjusting its vocal tones, gestures, and phrases to motivate, encourage, or correct users based on their responses. This adaptation created a more interactive and personalised experience for users. In contrast, the random robot's random behaviour left users with a less varied and unpredictable interaction. The majority of users' comments highlighted that the adaptive robot exhibited a greater range of animations, gestures, body movements, and feedback than the random robot, contributing to enhanced user engagement, expressiveness, and emotional relevance.

7.1.2 Thematic Analysis

We conducted a thematic analysis based on users' feedback gathered during the semi-structured interview [75]. We also considered some notes collected by test evaluators during the interactive sessions with the robot (i.e. users' comments during interactions). Three main themes were identified: one referred to engagement and stimulation, the attribution of human traits to robots, and the third referred to the robot's adaptions (or lack of) and coherence to the user responses and actions. The first theme help us to address RQ1 and RQ2.

Engagement and stimulation. The topic concerns the ability of robots to engage users and provide stimulation during interactions. Users observed and noted variances in the robots' capacity to actively involve them and effectively stimulate their participation during the interaction. In terms of user engagement and stimulation, 23 participants observed differences between the two robots. Specifically, they highlighted that the first robot exhibited greater involvement and encouragement, resulting in higher comfort during play. In particular, two participants (ID1 and ID16) emphasized that the adaptive robot's vocal feedback motivated them to improve and persist, not only when providing correct answers but especially when giving incorrect ones. ID1 reported: *"The robot would encourage me by saying, 'Come on, try again; we can do it.' It gave me the confidence to continue even when I was wrong."* Moreover, participants also noted differences in the language used, with the adaptive robot being more engaging than the random, sometimes providing brief responses. These adaptive behaviours supported several users in concentrating more on the task and maintaining a high focus and attention. ID 17 said: *"The adaptive robot stimulated me to stay focused and attentive, especially when it provided the answers."* These stimulating behaviours, coupled with animations that aligned with the message conveyed, contributed to the adaptive robot's expansive, encouraging, and stimulating behaviour" (as described by ID1). Furthermore, the adaptive robot's movements, such as approaching the user and displaying patience, reinforced the stimulation message, instilling a sense of reassurance and further engaging the participants. These heightened stimulation behaviours rendered the adaptive robot more helpful, engaging, and capable of putting users at ease. As ID1 articulated, *"The first (adaptive) robot made an effort to make me feel more comfortable when I made mistakes compared to the others."* Several participants acknowledged that this type of behaviour and attitude could significantly enhance attention and alleviate feelings of judgment among individuals who may be socially isolated. On the other hand, the feedback provided by the random robot seemed to bother some users. Some participants expressed frustration at the robot because it frequently reminded them to pay attention and focus more. The repetitive nature of this feedback amplified the significance of their errors. Although the adaptive robot also uttered similar feedback, participants noted that the adaptive robot delivered it in a more consistent and challenging manner. Overall, the differences in involvement and stimulation between the two robots considerably impacted user experiences and perceptions. The adaptive robot's engaging and encouraging behaviours created a more supportive and motivating environment. In contrast, the random robot's detached and repetitive feedback proved less effective and occasionally bothersome to the users.

Attribution of Human-Traits. This theme encompasses the expressiveness and human-like behaviour exhibited by the robot. Various users take note of the robots' capability to express emotions, engage in human-like gestures, and adopt behaviours reminiscent of human interactions. The extent to which the robots successfully convey a sense of humanity during the interaction is reported by users. During the interaction with both versions of the robot, several users ascribed humanized traits to the robots. These attributions were mainly based on the robot's animations, trajectories, and consistent speech in the adaptive version. Users mentioned that the adaptive robot's behaviour exhibited social, friendly, and helpful characteristics. Notably, several users observed that the adaptive robot desired to converse with individuals by adopting gestural behaviours similar to those commonly observed in Italian people. The robot's consistent behaviours conveyed the impression that it sought to contact the user, almost as if it aimed to establish a relationship (ID6 said: *"it looked at me as seemingly seeking contact, it seemed interested like a human being. Sometimes, I forgot that it was a robot, and...it is very human-like in comporting, especially in the second case" [which was the adaptive robot condition])*. The unexpected human-like behaviour exhibited by the robot was seen as endearing (ID8) because it deviated from 'typical' robotic behaviour. Its 'cooperative' attitude also contributed to a perception of the robot as possessing human-like characteristics. Additionally, various users compared their interactions with the robot to conversing with a child. After the interaction, users expressed gratitude towards the robot and gently patted it on the head, treating it like a human. Some users asked if the robot wanted to accompany them home. A few users reported temporarily forgetting that they were interacting with a robot. In summary, most users attributed human-like traits to the adaptive robot based on their observed behaviour during the interaction. The robots' animations, trajectories, and consistent speech in the adaptive version significantly elicited these perceptions. Users treated the robots as humans, expressing gratitude and physical gestures and engaging in conversations resembling human-to-human interactions.

Adaptation and responsiveness. The last theme helps us to address RQ1. The behaviour of the adaptive robot was widely recognized by multiple users as highly encouraging, reassuring, and engaging, regardless of the robot's personality. Several users perceived the adaptive robot as responsive and adaptive to their "status". One user (ID2) expressed, *" It conveyed a lot of confidence to me, and in my opinion, it was adaptive to my responses, which reassured me."* The synchronization between animations and speech was highly appreciated by most users, as it resulted in more consistent, natural behaviour and a positive, kind attitude. As one user (ID 8) stated, *" In addition to encouraging me to continue the game, what it told me was perfectly in sync with the animations, creating a human-like behaviour. This synchronization created a combination of feedback and immersive animations."* This synchronicity generated a more pleasant, gentle, calm, and cooperative behaviour, as if the robot was attempting to communicate in a more present and human-like manner. Moreover, this consistency in behaviour gave several users the impression of being with a coach who stimulated and motivated them to improve continuously while instilling confidence and peace of mind. Some users even felt connected with the adaptive robot through its verbal content and movements. This connection was described as having a substantial positive impact on users, making them feel like conversing with a human. Furthermore, various users observed changes in the tone and speed of the adaptive robot's speech, emphasizing that these variations seemed linked to the game performances. For instance, one user (ID11) noted that the robot's speech would slow down and lower in tone when the user was less attentive while becoming more cheerful during more involved moments. In contrast, the random robot's inconsistent behaviour in terms of animations and feedback was generally disliked by the majority of users. This inconsistency conveyed a sense of inattentiveness, detachment, coldness, and occasionally theatrical and exaggerated behaviour. The random robot seemed uninvolved, giving users the impression that they were going to play alone. Some users particularly highlighted the chaotic nature of the random robot's behaviour, which led to a perception of unpredictability and an overall lack of naturalness. Specifically, one user(ID22) described the behaviour as "jerky, mechanical, and nervous," followed by moments of overly exaggerated sentences lacking coherence and appearing unnatural compared to the behaviour of the adaptive robot. In summary, the random robot was generally characterized as distant, disinterested, incoherent, and detached, which made users uncomfortable due to the unpredictability of its behaviour. The inconsistent behaviour of the animations and feedback offered by the random robot is something that most users did not like.

7.2 RQ2. Does adaptive robot behaviour have an influence on user engagement and user experience?

To answer RQ2, we used the thematic analysis described before, the UES-short questionnaire to assess the impact on engagement, and four statements to assess likeability and satisfaction, which are important UX dimensions. Regarding the likeability of robot interaction, three statements were evaluated: "I liked the responses (vocal feedback) that this robot gave me", "I liked the animations performed by this robot," and "I enjoyed interacting with this robot." The normality of the data was assessed for the statements, and the first two were found to be normally distributed. Specifically, the first statement exhibited data normality ($W = 0.92, p = 0.06$), while the second statement had data normality ($W = 0.92, p = 0.0556$). Subsequently, paired t-tests were conducted, revealing significant differences between the two conditions. The adaptive condition received more positive scores than the random robot regarding the provided verbal feedback with ($t = -3.65, p = 0.0010$). Similarly, the animations of the adaptive condition were favoured more than those of the random robot with ($t = -2.19, p = 0.0036$). However, for the third statement, "I enjoyed interacting with this robot," the data deviated from normality ($W = 0.83, p = 0.0004$). Consequently, a paired samples Wilcoxon test was employed, and the results ($V = 5.5, p = 0.0022$) indicated a significant difference between the two conditions. Overall, users preferred interacting with the adaptive robot (M=4.57, SD=0.97) over the random robot (M=3.96, SD=0.85). Lastly, the statement "I felt comfortable interacting with this robot" did not exhibit significant differences between the conditions, as evidenced by the Wilcoxon test ($V = 31, p = 0.16$).
Additionally, we evaluate the UES-short questionnaire [76]. Table 11 displays the robot's UES scores during the first and second interactions for each condition. Internal reliability of the user response was evaluated by calculating Cronbach's alpha coefficient for both interactions ($1st$ interaction $\alpha = 0.91$, $2nd$ interaction $\alpha = 0.92$. An $\alpha > 0.9$ = excellent reliability level [77]). The scale that yielded better results for both robot conditions during the first interaction was the Perceived Usability (PU).
In the first interaction, the adaptive robot received a higher mean score for the perceived usability (PU) scale (M=4.69, SD=0.52) compared to the random robot (M=4.5, SD=0.75). This indicates that participants found the adaptive robot more user-friendly and easier to interact with, reflecting positively on the overall user experience. Additionally, the adaptive robot achieved a higher mean score for the focus attention (FA) scale (M=4.41, SD=0.78) compared to the random robot (M=4.2, SD=1.04).

Table 11: UES values that were collected after the first and second interaction with the robot's adaptive and random conditions (1-5 scale is used in which 1= strongly disagree, and 5= strongly agree).

First Interaction		
Scale	M(SD) Adaptive Robot	M(SD) Random Robot
FA	4.41 (0.78)	4.2(1.04)
PU	**4.69**(0.52)	**4.5**(0.75)
AE	4.07(0.88)	4(0.87)
RW	4.45 (0.66)	4.38(1.02)
Overall Score	**4.40**	4.32
Second Interaction		
Scale	M(SD) Adaptive Robot	M(SD) Random Robot
FA	**4.65**(0.71)	4.11(0.98)
PU	4.57(0.82)	**4.40**(1.02)
AE	4.1(0.57)	3.88(1.13)
RW	4.73(0.61)	4.21(0.88)
Overall Score	**4.50**	4.15

During the second interaction, similar trends were observed. The adaptive robot received higher mean scores for the Focused Attention (FA) scale (M=4.65, SD=0.71) and the PU scale (M=4.57, SD=0.82) than the random robot.

Furthermore, when evaluating the overall scores for each interaction, the adaptive robot consistently outperformed the random robot. In the first interaction, the adaptive robot achieved an overall score of 4.40, slightly higher than the random robot's score of 4.32. While, in the second interaction, the adaptive robot obtained an overall score of 4.50, surpassing the random robot's score of 4.15. These results indicate that participants had a more positive overall experience with the adaptive robot than the random robot after trying both conditions (see Table 11).

7.2.1 Preferences

For the following analysis, the data of the preferences question and open questions in Appendix B are used.
86% of users preferred the interaction with the adaptation condition, compared to the random condition (14%). Then, we analyse the open questions we asked to motivate their preferences. The analysis employed a qualitative approach, categorizing user comments based on the reasons for preferring the adaptive or random robot. The analysis of user comments indicates a preference for the robot with adaptive behaviour for several reasons. Users found the adaptive robot more engaging, noting that its responsive feedback during gameplay was more motivating than random responses. Users also appreciated its realism, characterised by relaxed movements and natural speech patterns, which felt less theatrical and more genuine. Many users highlighted the robot's coherent and proactive engagement, especially its supportive actions during mistakes. The timely and motivational feedback were key distinguishing features. Overall, the adaptive robot was preferred for its likability, liveliness, and smoother, more enjoyable interaction experience. In summary, users favoured the robot with adaptive behaviours due to its perceived higher realism, stimulation, likeability, fluid

Table 12: Godspeed results for Adaptive (A) and Random robot (N)

Category	M(SD) Adaptive Robot	Random Robot	Significant
Anthropomorphism	**3.6** (1.04)	2.90 (1.23)	* $p < 0.05$
Likeability	**4.50** (0.63)	3.94 (1.07)	* $p < 0.05$
Perceived Intelligence	**3.89**(0.94)	3.6(0.92)	ns

interaction, and coherent feedback and animations. They appreciated the adaptive robot's gestures, vocal tone variations, and encouraging phrases.

7.3 RQ3. Does the adaptive robot behaviour affect user perception in terms of likeability, perceived intelligence, and anthropomorphism?

Table 12 presents the outcomes of administering the Godspeed questionnaire [74]. Specifically, the questionnaire assessed for both robot conditions' anthropomorphism, likeability, and perceived intelligence. Statistical tests were conducted to evaluate the presence of statistically significant disparities, employing the obtained scores from these categories of the two robot conditions. Concerning the anthropomorphism category, the normality of the data was assessed using the Shapiro-Wilks test, resulting in a homogeneous distribution (W = 0.97987, p-value = 0.8473). Subsequently, a Paired t-test was employed to determine whether a statistical difference existed regarding the anthropomorphic perception between the adaptive and the random condition. The analysis indicated a statistically significant difference between the two conditions (t = -2.5801, p-value = 0.01564). Participants perceived the adaptive robot to exhibit more anthropomorphism than the random condition. Concurrently, likeability and perceived intelligence were also evaluated. Both demonstrated a distribution that adhered to normality, with the following results from the Shapiro-Wilks test: likeability (W = 0.95488, p-value = 0.2621) and perceived intelligence (W = 0.98542, p-value = 0.9544). Consequently, a paired t-test was conducted to analyse potential distinctions between the adaptive and random robots regarding likeability and perceived intelligence. The analysis revealed that participants strongly preferred the likeability of the adaptive robot, as evident from the obtained results (t = -2.4008, p-value = 0.02352). No statistically significant differences were detected within the Perceived Intelligence category. Nevertheless, it is noteworthy that participants perceived the adaptive robot to possess greater intelligence than the random robot, with an average rating of 3.89.

8 Discussion

This study aimed to investigate the impact of robot adaptation behaviour, using a strategy trained with QL. The study sought to assess how the adaptive behaviour of the robot influenced user engagement, user perception (likeability, perceived intelligence, and anthropomorphism), and overall user experience. The quantitative and thematic analysis findings shed light on users' observations, feedback, and preferences related to the adaptive and random robot conditions. The participants perceived the interaction with the adaptive condition to be significantly more effective in terms of animations compared to the random robot (as determined by the paired t-test, with $p = 0.00011$). Users noticed changes in the adaptive robot's movements and animations, mainly when they provided consecutive correct answers or made consecutive errors. These observations highlight the adaptive robot's responsiveness, which contributed to a more immersive and interactive experience for users. In contrast, users interacting with the random condition did not notice significant changes in animations or gestures, perceiving its behaviour as random and lacking coherence. Regarding verbal feedback, the paired t-test results revealed, with a significant result, that users perceived the adaptive robot as providing more adaptive and stimulating verbal feedback than the random condition. Users reported noticing changes in the adaptive robot's speech patterns, such as tone and speech rate variations, based on their responses. This adaptation in verbal feedback contributed to a more engaging and motivating user interaction. On the other hand, users interacting with the random robot found its speech and behaviour random and sometimes incoherent, leading to a less predictable interaction. These findings support the hypothesis that users are sensitive to robot adaptive behaviours, and they can differentiate between them. The thematic analysis provided deeper insights into users' experiences and perspectives regarding the adaptive and random robot conditions. The themes identified included engagement and stimulation, attribution of human traits to robots, and adaptation and responsiveness. Users noted that the adaptive robot exhibited greater involvement, encouragement, and stimulation, resulting in higher comfort and motivation during the interaction. This finding helps us to address RQ1 and RQ2. Then we evaluate the influence of adaptive robot behaviour on user engagement and user experience. The study utilized the UES-short questionnaire to assess the impact on engagement and evaluated likability and satisfaction as important dimensions of the user experience. The UES scores obtained during the first and second interactions demonstrated higher overall engagement with the adaptive robot. The perceived usability (PU) scale yielded better results for the adaptive robot conditions, indicating that users found the adaptive robot more usable. The adaptive robot received higher scores in terms of engagement, usability, focus attention, and overall user experience. A majority of users (86%) expressed a preference for more interaction with the adaptive robot, as opposed to the random condition. The qualitative analysis of user comments revealed several reasons behind this preference. Users appreciated the adaptive robot's engagement and responsiveness to their actions. The adaptive robot's realism, natural speech rhythm, and coherent feedback were seen as distinguishing factors. Users found the adaptive robot more likeable, lively, and engaged in movements and vocal interaction. Finally, we examine the influence of adaptive robot behaviour on user perception, specifically in terms of likeability, anthropomorphism, and perceived intelligence, including adaptive animations, verbal feedback, and synchronization between animations and speech, which were perceived as more engaging, supportive and human-like anthropomorphism. Users found the adaptive robot more likeable and attributed more anthropomorphism than the random robot. These results suggest that the adaptive robot's ability to adapt its behaviour and engage with users in a more human-like manner influenced users' perceptions and preferences. Adaptive behaviours were perceived as more engaging, supportive, and human-like, including adaptive animations, verbal feedback, and synchronization between animations and speech. These findings have implications for designing and developing socially interactive robots, emphasizing the importance of incorporating adaptive behaviours to enhance user experiences and foster more meaningful interactions.

9 Conclusion

In conclusion, we designed and defined parameters to support an intelligent algorithm to manipulate the robot's behaviour adaptation in a simulation and real user test. With the help of a psychotherapist, we identified three typologies of actions that can positively affect the user during a serious game scenario. We trained a QL algorithm to guide the robot by identifying the most suitable action based on the user's state. The training involved a simulation in which the agent interacted, one at a time, with two user profiles, one representing a user with MCI and one representing a healthy user while playing a serious game. After training the algorithm, we performed a user study to evaluate the perceived adaptation of the robot with 28 real users. During the test, users interacted with a robot that followed the adaptive behaviour, trained in the simulation, and a robot performing random behaviour. The results showed that users recognised the two robot behaviours and found the adaptive behaviour more interactive, stimulating, and consistent with human-like behaviour. However, it is important to recognise some limitations of the study. The parameters used in the simulation were extrapolated with the knowledge of previous experiments and with the help of experts with in-depth knowledge of the MCI target. The definition of these parameters and the simplification of the user states were done to simulate a cognitively impaired user and not in a real environment. However, this simplification only partially represents an authentic user interaction because other variables are not considered. In the future, we will research how the simulation can be improved towards this direction. In addition, the sample size was relatively small, and the study focused on a specific task and interaction scenario. Future research could expand the sample size and recruit older adults with MCI. Likewise, this study may suffer from the so-called novelty effect, so a more extended study should explore users' perceptions over time. Overall, the discussion of the findings highlights the significance of adaptive behaviours in robots and their impact on user perception, engagement, and stimulation.

Acknowledgement

This work is partly supported by the CNR project SERENI.

Ethical Approvals

The study was approved by the Regional Ethics Committee for Clinical Trials of the Tuscany Region, Section: Area Vasta Nord Ovest. The protocol was assigned the identifier GR-2019-12370776 (amendment approved on October 11, 2021).

Declaration of Interest

We declare that there are no conflicts of interest regarding the publication of this article. The authors have no financial or personal relationships with other people or organizations that could inappropriately influence or bias the content of this work.

References

[1] Cuayáhuitl, H., Frommberger, L., Dethlefs, N., Raux, A., Marge, M., Zender, H.: Introduction to the special issue on machine learning for multiple modalities in interactive systems and robots. ACM Transactions on Interactive Intelligent Systems (TiiS) **4**(3), 1–6 (2014)

[2] Tapus, A., Ţăpuş, C., Matarić, M.J.: User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. Intelligent Service Robotics **1**(2), 169–183 (2008)

[3] Yuan, F., Sadovnik, A., Zhang, R., Casenhiser, D., Paek, E.J., Zhao, X.: A simulated experiment to explore robotic dialogue strategies for people with dementia. Journal of Rehabilitation and Assistive Technologies Engineering **9**, 20556683221105768 (2022)

[4] Fong, T., Nourbakhsh, I., Dautenhahn, K.: A survey of socially interactive robots. Robotics and autonomous systems **42**(3-4), 143–166 (2003)

[5] Akalin, N., Loutfi, A.: Reinforcement learning approaches in social robotics. Sensors **21**(4), 1292 (2021)

[6] Hemminghaus, J., Kopp, S.: Towards adaptive social behavior generation for assistive robots using reinforcement learning. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp. 332–340 (2017)

[7] Keizer, S., Ellen Foster, M., Wang, Z., Lemon, O.: Machine learning for social multiparty human–robot interaction. ACM transactions on interactive intelligent systems (TIIS) **4**(3), 1–32 (2014)

[8] De Greeff, J., Belpaeme, T.: Why robots should be social: Enhancing machine learning through social human-robot interaction. PLoS one **10**(9), 0138061 (2015)

[9] Tsiakas, K., Dagioglou, M., Karkaletsis, V., Makedon, F.: Adaptive robot assisted therapy using interactive reinforcement learning. In: International Conference on Social Robotics, pp. 11–21 (2016). Springer

[10] Ritschel, H., Seiderer, A., Janowski, K., Wagner, S., André, E.: Adaptive linguistic style for an assistive robotic health companion based on explicit human feedback. In: Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments, pp. 247–255 (2019)

[11] Pou-Prom, C., Raimondo, S., Rudzicz, F.: A conversational robot for older adults with alzheimer's disease. ACM Transactions on Human-Robot Interaction (THRI) **9**(3), 1–25 (2020)

[12] Tapus, A.: Improving the quality of life of people with dementia through the use of socially assistive robots. In: 2009 Advanced Technologies for Enhanced Quality

of Life, pp. 81–86 (2009). IEEE

[13] Magyar, J., Kobayashi, M., Nishio, S., Sinčák, P., Ishiguro, H.: Autonomous robotic dialogue system with reinforcement learning for elderlies with dementia. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 3416–3421 (2019). IEEE

[14] Modares, H., Ranatunga, I., Lewis, F.L., Popa, D.O.: Optimized assistive human–robot interaction using reinforcement learning. IEEE transactions on cybernetics **46**(3), 655–667 (2015)

[15] Lee, K.M., Peng, W., Jin, S.-A., Yan, C.: Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. Journal of communication **56**(4), 754–772 (2006)

[16] Dryer, D.C.: Getting personal with computers: how to design personalities for agents. Applied artificial intelligence **13**(3), 273–295 (1999)

[17] Aronovitch, C.D.: The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. The Journal of social psychology **99**(2), 207–220 (1976)

[18] Shibata, T.: Therapeutic seal robot as biofeedback medical device: Qualitative and quantitative evaluations of robot therapy in dementia care. Proceedings of the IEEE **100**(8), 2527–2538 (2012)

[19] McColl, D., Nejat, G.: Meal-time with a socially assistive robot and older adults at a long-term care facility. Journal of Human-Robot Interaction **2**(1), 152–171 (2013)

[20] Pino, O., Palestra, G., Trevino, R., De Carolis, B.: The humanoid robot nao as trainer in a memory program for elderly people with mild cognitive impairment. International Journal of Social Robotics **12**, 21–33 (2020)

[21] Winkle, K., Caleb-Solly, P., Turton, A., Bremner, P.: Social robots for engagement in rehabilitative therapies: Design implications from a study with therapists. In: Proceedings of the 2018 Acm/ieee International Conference on Human-robot Interaction, pp. 289–297 (2018)

[22] Kubota, A., Riek, L.D.: Methods for robot behavior adaptation for cognitive neurorehabilitation. Annual review of control, robotics, and autonomous systems **5**, 109–135 (2022)

[23] Ferreira, E., Lefevre, F.: Reinforcement-learning based dialogue system for human–robot interactions with socially-inspired rewards. Computer Speech & Language **34**(1), 256–274 (2015)

[24] Leite, I., Pereira, A., Castellano, G., Mascarenhas, S., Martinho, C., Paiva, A.: Modelling empathy in social robotic companions. In: International Conference on User Modeling, Adaptation, and Personalization, pp. 135–147 (2012). Springer

[25] Gordon, G., Spaulding, S., Westlund, J.K., Lee, J.J., Plummer, L., Martinez, M., Das, M., Breazeal, C.: Affective personalization of a social robot tutor for children's second language skills. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 30 (2016)

[26] Ritschel, H., Baur, T., André, E.: Adapting a robot's linguistic style based on socially-aware reinforcement learning. In: 2017 26th Ieee International Symposium on Robot and Human Interactive Communication (ro-man), pp. 378–384 (2017). IEEE

[27] Tay, B., Jung, Y., Park, T.: When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. Computers in Human Behavior **38**, 75–84 (2014)

[28] Watkins, C.J., Dayan, P.: Q-learning. Machine learning **8**, 279–292 (1992)

[29] Breazeal, C.: Toward sociable robots. Robotics and autonomous systems **42**(3-4), 167–175 (2003)

[30] Duffy, B.R.: Anthropomorphism and the social robot. Robotics and autonomous systems **42**(3-4), 177–190 (2003)

[31] Severinson-Eklundh, K., Green, A., Hüttenrauch, H.: Social and collaborative aspects of interaction with a service robot. Robotics and Autonomous systems **42**(3-4), 223–234 (2003)

[32] Scherer, K.R.: Personality inference from voice quality: The loud voice of extroversion. European Journal of Social Psychology **8**(4), 467–487 (1978)

[33] Woods, S., Dautenhahn, K., Kaouri, C., Boekhorst, R., Koay, K.L., Walters, M.L.: Are robots like people?: Relationships between participant and robot personality traits in human–robot interaction studies. Interaction Studies **8**(2), 281–305 (2007)

[34] Looije, R., Neerincx, M.A., Cnossen, F.: Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. International Journal of Human-Computer Studies **68**(6), 386–397 (2010)

[35] Celiktutan, O., Gunes, H.: Computational analysis of human-robot interactions through first-person vision: Personality and interaction experience. In: 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 815–820 (2015). IEEE

[36] Manca, M., Paternò, F., Santoro, C., Zedda, E., Braschi, C., Franco, R., Sale, A.: The impact of serious games with humanoid robots on mild cognitive impairment older adults. International Journal of Human-Computer Studies **145**, 102509 (2021)

[37] Paradeda, R.B., Hashemian, M., Rodrigues, R.A., Paiva, A.: How facial expressions and small talk may influence trust in a robot. In: International Conference on Social Robotics, pp. 169–178 (2016). Springer

[38] Salichs, M.A., Encinar, I.P., Salichs, E., Castro-González, Á., Malfaz, M.: Study of scenarios and technical requirements of a social assistive robot for alzheimer's disease patients and their caregivers. International Journal of Social Robotics **8**(1), 85–102 (2016)

[39] Mead, R., Matarić, M.J.: Autonomous human–robot proxemics: socially aware navigation based on interaction potential. Autonomous Robots **41**(5), 1189–1201 (2017)

[40] Ellis, A., Abrams, M., Abrams, L.: Personality Theories: Critical Perspectives. Sage, ??? (2008)

[41] Robert, L.: Personality in the human robot interaction literature: A review and brief critique. In: Robert, LP (2018). Personality in the Human Robot Interaction Literature: A Review and Brief Critique, Proceedings of the 24th Americas Conference on Information Systems, Aug, pp. 16–18 (2018)

[42] Robert Jr, L.P., Alahmad, R., Esterwood, C., Kim, S., You, S., Zhang, Q., *et al.*: A review of personality in human–robot interactions. Foundations and Trends® in Information Systems **4**(2), 107–212 (2020)

[43] Esterwood, C., Essenmacher, K., Yang, H., Zeng, F., Robert, L.P.: A meta-analysis of human personality and robot acceptance in human-robot interaction. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1–18 (2021)

[44] Goldberg, L.R.: The structure of phenotypic personality traits. American psychologist **48**(1), 26 (1993)

[45] McCrae, R.R., Costa, P.T.: Personality in Adulthood: A Five-factor Theory Perspective. Guilford Press, ??? (2003)

[46] Nass, C., Lee, K.M.: Does computer-synthesized speech manifest personality? experimental tests of recognition, similarity-attraction, and consistency-attraction. Journal of experimental psychology: applied **7**(3), 171 (2001)

[47] Jensen-Campbell, L.A., Gleason, K.A., Adams, R., Malcolm, K.T.: Interpersonal conflict, agreeableness, and personality development. Journal of personality

71(6), 1059–1086 (2003)

[48] Lippa, R.A., Dietz, J.K.: The relation of gender, personality, and intelligence to judges' accuracy in judging strangers' personality from brief video segments. Journal of Nonverbal Behavior **24**(1), 25–43 (2000)

[49] Isbister, K., Nass, C.: Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. International journal of human-computer studies **53**(2), 251–267 (2000)

[50] Kim, S.-Y., Kim, J.-M., Stewart, R., Kang, H.-J., Kim, S.-W., Shin, I.-S., Park, M.-S., Cho, K.-H., Yoon, J.-S.: Influences of personality traits on quality of life after stroke. European neurology **69**(3), 185–192 (2013)

[51] Koh, J.S., Ko, H.J., Wang, S.-M., Cho, K.J., Kim, J.C., Lee, S.-J., Pae, C.-U., Serretti, A.: The association of personality trait on treatment outcomes in patients with chronic prostatitis/chronic pelvic pain syndrome: an exploratory study. Journal of psychosomatic research **76**(2), 127–133 (2014)

[52] Esteban, P.G., Bagheri, E., Elprama, S.A., Jewell, C.I., Cao, H.-L., De Beir, A., Jacobs, A., Vanderborght, B.: Should i be introvert or extrovert? a pairwise robot comparison assessing the perception of personality-based social robot behaviors. International Journal of Social Robotics **14**(1), 115–125 (2022)

[53] Ekman, P., Friesen, W.V., O'Sullivan, M., Scherer, K.: Relative importance of face, body, and speech in judgments of personality and affect. Journal of personality and social psychology **38**(2), 270 (1980)

[54] Pittam, J.: Voice in Social Interaction vol. 5. Sage, ??? (1994)

[55] Riggio, R.E., Friedman, H.S.: Impression formation: The role of expressive behavior. Journal of personality and social psychology **50**(2), 421 (1986)

[56] Moon, Y., Nass, C.: How "real" are computer personalities? psychological responses to personality types in human-computer interaction. Communication research **23**(6), 651–674 (1996)

[57] Hara, F., Kobayashi, H.: Use of face robot for human-computer communication. In: 1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century, vol. 2, pp. 1515–1520 (1995). IEEE

[58] Norman: The Design of Everyday Things. MIT Press, [New York] (2014)

[59] Zedda, E., Manca, M., Paternò, F., Santoro, C.: Older adults' user experience with introvert and extravert humanoid robot personalities. Universal Access in the Information Society, 1–17 (2023)

[60] Zedda, E., Manca, M., Paterno, F., Santoro, C.: Mci older adults' user experience

with introverted and extraverted humanoid robot personalities. In: Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter. CHItaly '23. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3605390.3605405 . https://doi.org/10.1145/3605390.3605405

[61] Cheek, J.M., Buss, A.H.: Shyness and sociability. Journal of personality and social psychology **41**(2), 330 (1981)

[62] Zedda, E., Manca, M., Paternò, F., *et al.*: Towards adaptation of humanoid robot behaviour in serious game scenarios using reinforcement learning. In: ALTRUIST@ ICSR, pp. 93–99 (2022)

[63] Zedda, E., Manca, M., Paterno, F., Santoro, C.: An adaptive behaviour-based strategy for sars interacting with older adults with mci during a serious game scenario. arXiv preprint arXiv:2305.01492 (2023)

[64] Tsiakas, K., Abujelala, M., Makedon, F.: Task engagement as personalization feedback for socially-assistive robots and cognitive training. Technologies **6**(2), 49 (2018)

[65] Khamassi, M., Velentzas, G., Tsitsimis, T., Tzafestas, C.: Active exploration and parameterized reinforcement learning applied to a simulated human-robot interaction task. In: 2017 First IEEE International Conference on Robotic Computing (IRC), pp. 28–35 (2017). IEEE

[66] Krsmanovic, F., Spencer, C., Jurafsky, D., Ng, A.Y.: Have we met? mdp based speaker id for robot dialogue. In: INTERSPEECH, pp. 461–464 (2006)

[67] Chen, H., Park, H.W., Breazeal, C.: Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. Computers & Education **150**, 103836 (2020)

[68] Park, H.W., Grover, I., Spaulding, S., Gomez, L., Breazeal, C.: A model-free affective reinforcement learning approach to personalization of an autonomous social robot companion for early literacy education. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 687–694 (2019)

[69] Kim, G.H., Jeon, S., Im, K., Kwon, H., Lee, B.H., Kim, G.Y., Jeong, H., Han, N.E., Seo, S.W., Cho, H., *et al.*: Structural brain changes after traditional and robot-assisted multi-domain cognitive training in community-dwelling healthy elderly. PloS one **10**(4), 0123251 (2015)

[70] McCallum, S., Boletsis, C.: Dementia games: A literature review of dementia-related serious games. In: International Conference on Serious Games Development and Applications, pp. 15–27 (2013). Springer

[71] Cohen-Mansfield, J., Dakheel-Ali, M., Marx, M.S.: Engagement in persons with

dementia: the concept and its measurement. The American journal of geriatric psychiatry **17**(4), 299–307 (2009)

[72] Bellman, R.: The theory of dynamic programming. Bulletin of the American Mathematical Society **60**(6), 503–515 (1954)

[73] Cruz-Sandoval, D., Morales-Tellez, A., Sandoval, E.B., Favela, J.: A social robot as therapy facilitator in interventions to deal with dementia-related behavioral symptoms. In: 2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 161–169 (2020). IEEE

[74] Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. International journal of social robotics **1**, 71–81 (2009)

[75] Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative research in psychology **3**(2), 77–101 (2006)

[76] O'Brien, H.L., Cairns, P., Hall, M.: A practical approach to measuring user engagement with the refined user engagement scale (ues) and new ues short form. International Journal of Human-Computer Studies **112**, 28–39 (2018)

[77] Tavakol, M., Dennick, R.: Making sense of cronbach's alpha. International journal of medical education **2**, 53 (2011)

Appendix A

Table with transaction probabilities used in the simulation.

In the following tables, the values in each cell denote the probability of modifying the engagement level when the corresponding action is taken in the given state. For example, for an MCI user in a state with a High engagement (Appendix Table A1, title: From High (H) to →), if an engaging action (Action0) is taken, then in 2 out of 10 cases the user engagement level will decrease to Low (non-attentive user).

Table A1: Engagement probabilities for MCI user.

MCI Simulation - From High (H) to →			
Engagement Level	Action0	Action1	Action 2
Low	0.2	0.2	0.1
Intermediate	0.3	0.4	0.4
High	0.5	0.4	0.5
From Intermediate (I) to →			
Low	0.2	0.1	0.2
Intermediate	0.4	0.5	0.3
High	0.4	0.4	0.5
From Low (L) to →			
Low	0.3	0.5	0.2
Intermediate	0.4	0.3	0.4
High	0.3	0.2	0.4

Table A2: Engagement probabilities for a healthy user.

Healthy Simulation - From High (H) to →			
Engagement Level	Action0	Action1	Action 2
Low	0.1	0.2	0.2
Intermediate	0.4	0.2	0.3
High	0.5	0.6	0.5
From Intermediate (I) to →			
Low	0.1	0.1	0.2
Intermediate	0.5	0.4	0.5
High	0.4	0.5	0.3
From Low (L) to →			
Low	0.1	0.3	0.1
Intermediate	0.4	0.3	0.2
High	0.5	0.4	0.7

Appendix B

We present the questionnaire, a semi-structured interview, open questions and statements for the adaptation evaluation and UX assessment. These components were previously outlined in Section 6.2. The test was conducted in December 2022 with 28 users in a laboratory.

Semi-structured Interview for Robot Adaptation Evaluation

1. Have you noticed any changes in the robot?
2. Where and when? What were the most striking ones?
3. Did you notice a change in the way he spoke? Which one?
4. Did you notice a change in the way it moved? Which one and why?
5. Did you enjoy interacting with these robots? Why?
6. Which aspects did you enjoy most and which least?
7. Did one of the two seem more adaptive to you? Did it adapt well, or did it bother you?

Questionnaire personalised for Evaluation Robot Adaptation

Statements

Table B3: Statements for adaptation test. 1= strongly disagree while 5= strongly agree

	Statement	Scale
1	I liked the responses (vocal feedback) that this robot gave me.	1-5
2	I liked the animations performed by this robot	1-5
3	I enjoyed interacting with this robot.	1-5
4	I felt comfortable interacting with this robot.	1-5
5	I noticed a change or increased stimulation in the way the robot gave me voice feedback (the robot responses)	1-5
6	I noticed some changes in the animations or movements performed by the robot.	1-5
Preference Question	Which robot do you prefer more?	1st or 2nd

Open questions

1. Why did you like or dislike the responses (vocal feedback) the robot gave?
2. Why did you like or dislike the animations the robot gave?
3. Give reasons why you liked or disliked interacting with this robot.
4. If you noticed any change, when, where, or what did the robot say to make you notice this change?
5. If you noticed any changes in the robot's movements or animations, when, where did you notice these changes?
6. Is there anything you would like to improve in speech or robot animations?