The Impact of Serious Games with Humanoid Robots on Mild Cognitive Impairment Older Adults

Marco Manca[*], Fabio Paternò[*1], Carmen Santoro[*], Eleonora Zedda[*], Chiara Braschi[+], Roberta Franco[+], Alessandro Sale[+]
[*]CNR-ISTI, HIIS Laboratory, Pisa, Italy
[+]CNR-IN, Pisa, Italy

ABSTRACT
The number of Mild Cognitive Impairment older adults is increasing; thus, it becomes more and more important to provide them with support to avoid cognitive decline. Interactive serious games can play an important role to this end. So far most of them have been deployed through tablets, but the emerging humanoid robots may open up new possibilities. We present a study aiming to better understand the impact of humanoid robots in supporting serious games for such users. For this purpose, two versions of a serious game have been designed by a multi-disciplinary team, one for humanoid robots and one for tablets. We report on its use by a group of such users, and discuss their experience.

Author Keywords
Mild cognitive impairment, serious games, interactive technologies, humanoid robots

INTRODUCTION

With a senior population that is foreseen to more than double by 2050 worldwide[2], an increasing demand for high-quality elderly support is likely to be expected in the coming years. Among the various disabilities typically associated with ageing, cognitive impairments are those affecting a significant part of people aged 65 plus. For instance, the Alzheimer Association reported that approximately 15 to 20 percent of people age 65 or older have Mild Cognitive Impairment (MCI).[3] MCI is an intermediate stage between the cognitive decline associated with normal ageing and more serious forms of dementia. Seniors with MCI often show memory loss or forgetfulness and may have issues with other cognitive functions such as language, attention and visuospatial abilities. The potential evolution of this disease makes it necessary to increasingly provide aid to such people over time, and it is frequently associated with an increased burden on their caregivers, and in the worst case, the institutionalization of such patients. Thus, it is especially important to provide people who are potentially at risk of developing dementia with timely and engaging cognitive training to slow the progression of their decline, while significantly cutting down the associated socio-economic costs.

Currently, cognitive training for seniors with a diagnosis of MCI is typically administered by professional caregivers who often use paper–based material, less frequently enhanced by computer-based support. While the introduction of computerised serious games for cognitive support has received several proposals (Chan et al., 2016; Savulich et al., 2017; Vaportzis et al., 2017), nevertheless the possibilities that existing interactive technologies may offer to such people seem largely untapped until now. The issues in adopting

---

[1] Corresponding author: Fabio Paternò, email:fabio.paterno@isti.cnr.it

[2] https://www.who.int/news-room/fact-sheets/detail/ageing-and-health, Retrieved September 19, 2019

[3] https://www.alz.org/alzheimers-dementia/what-is-dementia/related_conditions/mild-cognitive-impairment, Retrieved September 19, 2019

such technologies can be explained by the fact that, for seniors with MCI, technology not only can provide opportunities, but may also present potential challenges because of their difficulties in visual attention and memory amongst other cognitive abilities. Thus, to allow older adults with cognitive impairments to interact comfortably with technology and actually benefit from it, it is important to understand how they approach and perceive the cognitive interventions provided through various types of assistive interaction technologies. In this work, we focus on two kinds of interactive technologies: humanoid robots and tablets.

On the one hand, the tablet has been the interactive technology most considered till now, also due to its low cost. Some work (Joddrell and Astell, 2016) has suggested that even people affected by dementia may be able to use touch screen devices effectively, which may provide some benefits to them and their caregivers. On the other hand, we found it interesting to investigate the impact of the use of humanoid robots on interventions targeting people with cognitive disabilities. In this context, such technology has been less explored, although some initial attempt has been made in this area.

In this paper we report on a study whose goal was to investigate how seniors with mild cognitive impairment relate with and perceive serious games accessed through humanoid robots, as part of a training programme aimed to improve their cognitive status. To this goal, we designed and developed a quiz game to help the subjects to train their memory while receiving encouraging feedback from the application. We introduced not only the version of the game for humanoid robot but also a version for tablets, whose purpose was to represent a useful reference point for our analysis, given that the tablet is currently the most used device in cognitive training.

The structure of the paper is as follows. In the next section we analyse the state of the art, then we provide the context in which the described study was carried out, and identify relevant requirements for applications supporting cognitive stimulations targeting MCI seniors. In the next section we describe the music quiz game designed to support cognitive stimulation of seniors. We report on the user study in which 14 users used two different versions of the application (tablet-based and robot-based). Finally, we provide a discussion of lessons learnt and conclude with some remarks and indications for future work.

RELATED WORK

Mild Cognitive Impairment (MCI) is characterized by a cognitive decline greater than expected as compared to that typically associated with an individual's age, but which does not significantly interfere with daily activities (Gauthier et al., 2006). Due to the increasingly high numbers of the ageing people expected to suffer from cognitive impairments in the coming years (Prince et al., 2013), many interventions have been promoted to aid this population under a wide variety of perspectives. For instance, Mentis et al. (2019) discuss the security and privacy risks associated with online services (e.g. banking) to which seniors suffering from MCI are increasingly exposed nowadays. However, one of the areas that has received the most attention is related to interventions aiming to delay the cognitive decline of MCI population, which is at higher risk of developing dementia. A large body of literature has focused on the development of Serious Games (SG), digital applications aiming to go beyond mere entertainment, thereby help assessment, stimulation, treatment and rehabilitation of patients suffering from cognitive disorders (Robert et al., 2014; Manera et al., 2015). A literature review of studies on the use of SG in various neurodegenerative disorders, MCI and Alzheimer's Disease (AD) is reported in (McCallum and Boletsis, 2013), which emphasises that, while physical games (i.e. games promoting physical fitness) can positively affect several health areas of players with MCI and mild AD, such as balance, gait, and voluntary motor control, cognitive games can improve attention, memory, and visuo-spatial abilities.

The approaches that improve cognitive aspects include cognitive training and cognitive engagement. Cognitive training targets specific domains with the expectation that improvements observed in one domain will be potentially transferred to other ones. In contrast, cognitive engagement involves acquiring new skills, which may simultaneously train several cognitive abilities including executive function, reasoning, and memory (Vaportzis et al., 2017; Chan et al., 2016). While cognitive engagement potentially offers greater opportunities to produce broader benefits compared to more focused cognitive training, it has received limited attention due to the cost and complexity of assessing users for prolonged periods in experimentally controlled real-world environments. Indeed, the type of cognitive intervention that has shown most benefits is that which engages individuals in a focused, repetitive task, usually targeting the improvement of one specific cognitive ability (e.g. memory, executive functioning): this is also the kind of intervention we address in this work. However, cognitive training packages are typically repetitive, thus two major challenges are overcoming the incidence of dropout rates in patient groups and increasing their engagement during such activities (Cohen-Mansfield et al., 2009). Savulich et al. (2017) conducted a randomized controlled trial of cognitive training using an iPad-based memory game in 42 patients with amnestic MCI assigned to either the cognitive training group or the control one. The cognitive training group maintained high levels of enjoyment and motivation to continue playing, with self-confidence, self-rated memory ability, and episodic memory improvement over time. Differently from this work, they did not include robots in training.

A number of studies showed that elderly people with cognitive disorders may have problems in using many SG currently available. This is due to poor familiarity with game technology, and often derives from the fact that most SGs have been developed for entertainment (e.g., Nintendo Wii Fit, Big Brain Academy), and with a “typical healthy user” in mind. To understand this problem, Manera et al. (2015) investigated the acceptability of a tablet-based SG to stimulate executive functions and praxis. The results concerning game performance (e.g. time spent playing, number of errors), as well as the self-reported data, confirmed the acceptability of the game for patients with MCI, AD and related disorders, and the usefulness of employing it for training purposes.

Another emerging area for improving cognitive skills in older adults involves robots. Some preliminary work (Churamani et al., 2018) explored the use of emotion recognition within dialogues with the NAO robot, to make cognitive training more emotionally engaging for seniors. Other work explored whether robot-based SGs can be beneficial even for people with dementia. In particular, in (Tapus et al., 2009) and (Martìn et al., 2013) robots were exploited in combination with music-based games targeting people with dementia. In (Tapus et al., 2009) the robot engages users in a musical game, which adaptively adjusts its difficulty to the abilities of the player. Results show that this approach may engage patients and keep them interested in interacting with the robot. In (Martìn et al., 2013) the authors present an application for dementia patients, in which the behaviour of the humanoid robot and the therapy session are visually programmed in a script that allows music playing, physical movements, speech synthesis and the activation of lights/leds on the humanoid robot. Initial results of experiments with real institutionalized patients affected by moderate and severe dementia showed a slight improvement in neuropsychiatric symptoms compared to other traditional methods but without significant statistical differences with respect to a baseline group. Both of such works used music and humanoid robots to activate patients affected by dementia, whereas we focus on MCI patients.

Focusing on older adults with MCI, (Stogl et al., 2019) investigated the use of a mobile robot for motor stimulation of people with MCI. They used an omnidirectional robot with handlebars and a force-torque sensor to support interaction with users. Ten older adults with MCI evaluated the device and the training. The results show that, during training, users tend to be more precise and faster in controlling the device, which indicates an improvement of their motor skills. They also felt safe and managed to adapt to changes in the device’s behaviour, suggesting that this device can be suitable for training. While they mainly address patients' motor skills via non-humanoid robots, we aim to improve seniors' memory using social robots.

Humanoid robots seem more promising since they can support more engaging interactions with users. Pino et al. (2019) present an approach for slowing the progression of cognitive decline in MCI patients by using a humanoid robot that supports tasks from the usual memory-training program. Subjects either had the support of the NAO robot or only that of a psychologist. The resulting data indicated that memory training with NAO increased the patients' visual gaze and reinforced the therapeutic behaviour compared to the other condition. An aspect highlighted in that work was the users' positive reaction to reinforcement phrases provided by the robot as a feedback after completing a task, and the importance of providing such feedback in a personalised manner (i.e. by including the name of the user who replied correctly). However, this was rather problematic in that work: the interaction with the robot was carried out in groups, and NAO was expected to recognize people by looking at their faces, which did not work well. In our work, each user interacts with the Pepper robot individually: this can support more reliable opportunities for personalised training.

When planning cognitive interventions including robots, it is important to analyse the attitude of older adults towards this technology. (Wu et al., 2016) explore the needs of older adults with MCI and their attitudes toward an assistive robot. Although participants reported difficulties in managing some daily activities, they did not see themselves as needing an assisting robot, but they considered it potentially useful either for themselves in the future, or for others who are very old, frail, alone or lonely. The goal of (Bechade et al., 2019) was to find an objective procedure for evaluating dialogues with a social robot, by collecting data from end-users: they first collected data though a Wizard of Oz system with Nao, next through an autonomous system with the Pepper robot. Although the dialogues with robots were positively perceived by users, some potential participants declined to participate since they did not want to interact with a robot: this shows that seniors are not used to talking to devices, thus robots should be introduced gradually.

The integration of social robots in a smart environment is the subject of (Schroeter et al., 2013): a socially assistive robot companion, which supports reminders, recommendations, video contacts with relatives/friends, and a cognitive stimulation game. Six trails involving patients suffering from MCI, AD and Front-Temporal Dementia were conducted. The authors report that, although speech recognition turned out to be insufficiently accurate for intuitive use, subjects, despite some initial scepticism, became convinced by the idea of support by the combination of a smart home with a social robot. The robot was valued for its embodied interactions, its ability to give reminders/suggestions, the fact that it comes physically to the user, talks to her, shows initiative and has some 'personality'. While in that paper the authors' aim was mainly to get qualitative feedback on how a robot can assist people in their life, our study is more focussed, addressing cognitive, memory-related abilities of MCI people.

In particular, we discuss the impact of two technologies (tablet and humanoid robot) in supporting serious games for seniors with MCI, by reporting on a multidisciplinary work that involved a group of such users in testing a serious game that exploited musical quizzes to provide engaging cognitive training. To the best of our knowledge, it is the first time that tablets and social robots have been exploited in parallel for improving memory abilities in MCI seniors, who individually interacted with the developed application.

CONTEXT

This research is the result of a collaboration between two groups working in two Institutes of the same campus: one group belonging to an Informatics institute and working in the area of interactive technologies, and the other one working in the Neuroscience field. The neuroscience group participates in a local project called *Train the Brain* (TTB) aiming to help aged MCI people to keep active their mind and body, and to prevent and/or slow down cognitive decline. They recruited a number of patients at risk or already suffering from a slight cognitive deficit. All participants were firstly evaluated by memory-disorders specialists at the local university clinic, since the diagnosis of MCI was based on the assessment of relevant neurological and clinical tests/examinations, as well as of personal data (e.g. socio-demographic data, medical history, pharmacological drug use and lifestyle habits).

In the TTB programme participants use games, social activities, and physical activity with a series of increasingly complex exercises as well as meetings, music therapy, group stories, bike and stretching exercise in order to tone at the same time their body and their brain. The programme is structured into 8 cycles, each one is divided into 18 sessions of activities designed to stimulate various cognitive functions: auditory and visual attention, visual-spatial memory, imagination, space-time and personal orientation, verbal memory, lexical skills, and affective memory. Two 60-minute sessions a day are scheduled, three times a week. Each cycle lasts three weeks, then the same sessions are proposed again, increasing the difficulties. In past editions of such project the caregivers tried to introduce computers for cognitive training but it was a failure. They used desktop computers with touch screens but the older adults had no familiarity with them, did not know how to manage window-based user interfaces, and found it very difficult to perform the training exercises with such platform. Patients explicitly expressed their dislike to carrying out the training games through PCs.

Thus, it was decided to propose the use of different interactive technologies to investigate whether they could be more useful and stimulating for MCI older adults. The decision was to use a tablet and a humanoid robot. Since most users had experience with smartphones, the tablet seemed something more consistent with their experience with technologies and also because touch-based devices make less demand of hand-eye coordination when compared with a desktop computer using a mouse and cursor. The humanoid robot was something completely new for them, and it was deemed interesting to investigate the reactions it can stimulate in such audience. Thus, the two groups agreed to co-design a serious game for stimulating the older adults in some cognitive aspects, and implement it in two versions (tablet and robot).

DESIGN

Considering the previous experiences of the TTB psychologists with older adults, it was decided to design an application requiring users to recognise songs from the years when they were younger. The song recognition exercise is a task of retrograde memory (for known songs) and anterograde memory (for unknown songs) to which is potentially added also a component of autobiographical (retrograde) memory when the known songs evoke possible personal memories of the individuals.

The designed application plays the initial part of songs which were for the majority from the 50s and 60s, and the older adults have to select in a limited amount of time from a list of three potential answers provided by the application the singer or the song title. The choice of the "songs" stimuli was made on the basis of previous experience with these subjects, which showed that, in response to such kind of stimuli they reacted with interest, fun, and motivation. In addition, while other types of stimuli used in the past with MCI patients tended to involve only one specific cognitive resource, the song quiz proposed, although mainly aimed to stimulate memory, also indirectly involves other capabilities such as attention and even reasoning (i.e. stimulating deductive reasoning such as “it was a song from my youth, then the singer could be…”). Moreover, the subjects work in a program which also includes other kinds of exercises (e.g. gym), thus in a context of social exchange the music may provide positive emotions.

Various meetings with the psychologists of the TTB program were necessary for co-designing several aspects of the game for the two devices. In such meetings it was decided that each question must show three options to choose from: the correct answer, one that can be misleading (i.e. it can be mistaken for the correct one because it has some similarities), and one that can be easily recognized as wrong.

Users can listen to a small segment of a song (20 seconds), and they should guess the title or the singer. In order to maintain a high attention level and not tire the older adults, we decided that each exercise would have proposed 15 songs. After the question is displayed and vocally synthesized by the application, a musical piece is played, starting from the beginning, and the users are asked to recognize the title or the singer. They have 20 seconds to respond, then the correct answer is shown: after this, the user can decide when to continue the game, by pressing a button in the UI: this was done to let users have the time to memorise the correct answer, especially when they wrongly replied to a question. When the users respond, if they indicate a wrong answer, the right one is shown by highlighting it in green, and the wrong answer is presented by using a red background. No response after 20 seconds is considered an error. The application shows the answer either after the user replies or when the time available is over.

The songs were provided by the psychologists. In order to stimulate memory tasks, they chose songs belonging to the decades from 1920 to 1970. In this way each user may recognise some childhood songs, others songs that may have been listened to at a later age, and other more modern songs. In total, 80 songs were identified. From this list, 15 were selected for the game sessions. In each game session such songs were randomly proposed. The three possible answers for each of them were always the same but presented in random order, to avoid using spatial memory to decide the answer.

The interaction was designed to take place in a multimodal way, also accounting for the behaviour of the older adults (i.e. it provides different feedback, depending on whether the user answers correctly or not). The two versions differ in the way users interact with them, as the robot also uses body movements, coloured eye and ear leds, its synthesised voice, and hand and arm gestures while speaking.

For the graphical part, we considered guidelines to design interfaces for elderly users (Hoffman and Hancock, 2015); W3C, 2019; Johnson and Finn, 2017; De Almeida, Ferreira and Soares, 2015; Park, Goh and So, 2014), such as use clear icons, use simple and common words, minimize the keyboard use, maximize contrast, fonts should be bigger than 12 pt, place icons near labels in order to reinforce the concept, use few colours, button should be large.

Finally, the system was enhanced in such a way to log significant information associated with user interactions: it was decided to record data regarding users' reaction times for answering each question, the time needed to complete the entire game in each session, the number of correct responses, the number of incorrect responses, the number of answers not provided within the time interval of 20 seconds, and the number of sessions needed by users to respond correctly to (at least) 12 questions (which correspond to 80% of total number of questions in each game session: 15). This last measure was considered by the psychologists as providing indications about users' learning rate.

THE GAME

We identified four main states for the Music Quiz game: login, menu, play and results. The starting state is login, where users access the game. The next state is "menu", where an introduction vocally presents the instructions. At the beginning of the introduction users are greeted by using their first names, to make them perceive the communication as more personalised. From the "menu" state the user can go to the "result" or the "play" state. After the users play the game, they enter into the 'result' state, where the results of the game are shown (number of right/wrong answers, mean reaction times and total session time).

The "play" state is divided into two sub-states ("question" and "answer"). The "question" state corresponds to displaying and vocally synthesizing the question, a progress bar with the remaining time to answer the question, and the three possible answers proposed in the form of large buttons, disabled at the beginning. After the question has been completely presented, the song and the countdown start, and the possible answers are enabled.

The user interface associated with the "answer" state changes depending on the answer chosen by the user. On the one hand, if the answer is wrong, it highlights the choice in red and displays the thumbs down icon in order to reinforce the negative feedback (see Figure 1b); it also shows the right answer (to allow users to learn it), with a green background and a thumbs up icon (see Figure 1c). On the other hand, if the user provides the right answer, a reinforcement sentence is shown (i.e. "very good, you gave the right answer"), and the singer and the song title are shown with a green background and thumbs up icon.

After the user provides the answer, for each type of answer (right, wrong or timeout) there are some possible feedbacks, randomly chosen from time to time. The possible feedbacks for correct answers are: "Congratulations! Correct answer" or "Very good! Right answer". The feedbacks for the wrong answers can be: "I'm sorry, you gave the wrong answer" or "Error". In case of timeout, the provided feedback is "Sorry. Time is up".

While all the above-mentioned sentences are also vocally synthesized in the two considered cases, one difference between the two devices is the additional feedback that is provided by the robot after a user provide the answer to a specific question. Indeed, the humanoid robot also provides feedback in terms of additional modalities such as head, body, hand and arm movements, as well as sounds and coloured leds positioned on shoulders and eyes (which become green in case of right answers, red for mistakes, and blue before answers). In particular, in case of a question correctly answered the (positive) feedback provided by Pepper is one of the following three "happy" animations:

- Pepper rises both arms as to greet the user, the eye leds become yellow, and it also produces a sound that expresses joy (see Figure 2, left part);
- It rises the right arm with a closed fist, its eye leds are green and a "hey!" expression is vocally provided;
- It nods its head, and the leds on the eyes and the shoulders become green.

Either in case of a question incorrectly answered or a timed-out question, the feedback provided by Pepper is one of these three animations:

- This animation is composed of two gestures: It shakes the head and it raises both arms up to the level of the torso;
- It groans and moves the arms up and down, the leds on the eyes and the shoulders become highlighted (see Figure 2, right part);
- It shakes the head and vocally rendering an "oh!" expression, while the eye leds become yellow and red.

When the number of correct answers exceeds 12 (which is the 80% of the total number of questions) one of the following animations is chosen:

- Pepper raises the right arm twice, eye leds become yellow and the sound of a celebrating crowd is rendered;

- It raises both arms to simulate a "win" gesture, the eye leds become yellow and the same sound used in the previous animation is rendered.

Such animations were identified so as to make the robot provide expressions that to some extent resemble human-like ones, so that users interpret the robot's behavior as empathic as possible (according to user's current game result). We chose not to use negative animations in this context, following the feedback suggestions of psychologists to encourage the elderly to do increasingly better.

The game starts the next song only after explicit user request, in this way users have time to memorize the right answer.

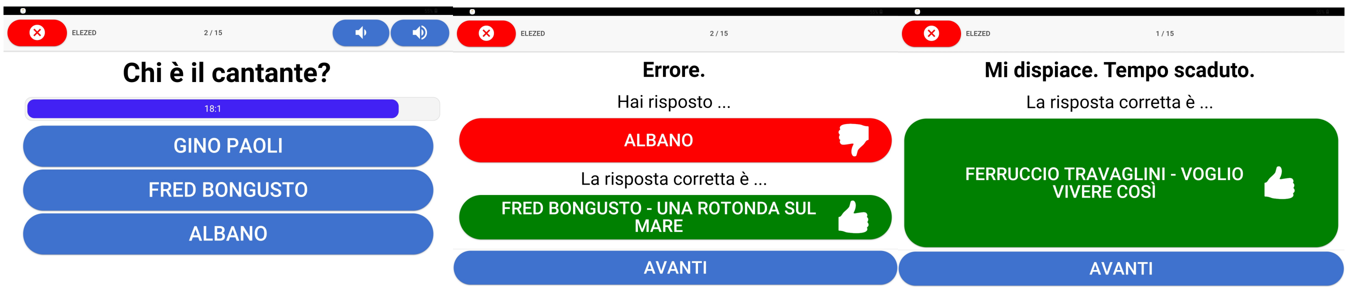

Figure 1: The application presenting a question (a), providing feedback about an incorrect answer (b), and about an occurred timeout (c).

Users can end the game at any time. At the end of a session a feedback message is displayed and synthesized depending on the percentage of correct answers:

- If the game is interrupted before the end, the message "See you next time" is presented.
- If the percentage of right answers is less than 50% the following sentence is presented: "You did a good job. See you next time to do even better!";
- If it is between 50% and 80% the sentence is: "Congratulations, you responded well to more than half of the stimuli. See you next time!";
- If the percentage is higher than 80% the phrase is: "Exceptional, great job! Keep it up!"

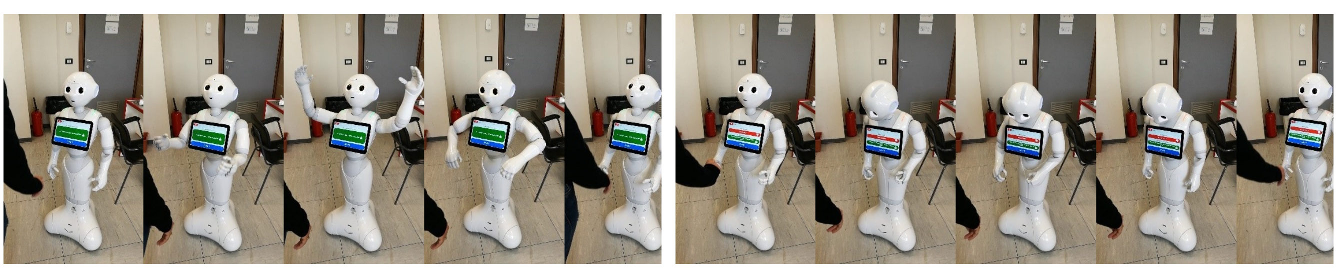

Figure 2: Example of positive (left part) and negative feedback (right part) from the robot

The tablet used for the experiment was a Samsung Galaxy Tab A SM-T580, with Android 8.1. The robot was a Pepper, by SoftBank Robotics, a 1.2-m-tall wheeled humanoid robot, with 20 degrees of freedom for motion in the body (17 joints for body language) and three omnidirectional navigation wheels to move around smoothly. It has a range of sensors to allow it to perceive objects and humans in its surroundings.

NAOqi OS is the robot ‘s operating system. It is a GNU/Linux distribution based on Gentoo, developed specifically to meet the needs of SoftBank Robotics robots. It provides and manages a number of programs and libraries required to control the robot sensors and movements. The NAOqi framework supports different programming languages (Java, Phyton, C++) and we decided to use Python because it is the most supported, and it can run both on PCs (for debugging purposes) and on the robot.

The graphical part of the game has been implemented with the Ionic framework. The part of the applications running on the robot tablet is a web application (HTML5 + JS + CSS) generated by Ionic and within the JavaScript part it is possible to raise events that can then be caught by the Python back-end module, which then calls the native functionalities of the robot (text to speech, led management and animations). The MusicQuiz Python module manages the robot during the execution of the games. The main goal of this component is to wait for events from the graphical user interface and then call the native functions which control the robot. The application running on the tablet is a native Android application generated by Ionic as well.

THE STUDY

The study was organised in twelve sessions conducted in the Train the Brain clinic during May-June 2019. Our participants were recruited by the local Train the Brain project, and all of them have had a diagnosis of Mild Cognitive Impairment. All participants were invited to provide their informed written consent before the beginning of the study. For each participant, demographic, and computer experience data were formally collected in a questionnaire before the first session.

The users were enrolled according to the following inclusion criteria: diagnosis of MCI, age over 65 years, Italian speaking participants.

Participants

14 participants (9 females) with age ranging between 69 and 84 (mean=75.3; std. dev=4.5) were involved in the experiment.

The users have different levels of education: 4 users held an elementary school degree, 2 middle school degree, 7 a high school degree, 1 an academic degree.

The level of computer experience was defined using four categories: 1) Very low knowledge on computer experience: very low experience with technologies and devices (call or sending messages), 2) Low knowledge: know the basic functionalities of devices (sending message and call, have internet connection), 3) Medium knowledge: able to use a smart device, 4) Good knowledge: able to use the functionality of smart device (navigate on the internet, using social and application). The elderlies have different levels of computer experience: five with very low knowledge, two with low knowledge, three with medium knowledge, four with good knowledge. The devices most used by them were mobile phones (all of them), then TVs, then computers (either desktop or portable ones). Only two users declared to use tablets (the type of device least used by the involved subjects).

Test Organisation

The experiment was organized in two different stages: 1) familiarizing with the humanoid robot Pepper and the tablet; 2) playing the music game with the humanoid robot and the tablet.

In the first phase participants received a brief introduction to the study, the devices used, and the main goals and motivations. Then, the participants saw for the first time the robot and could interact with it and the tablet in order to try the main game functionalities and familiarize themselves with them. While both devices were available when the study was introduced, it was interesting to note that all the participants were attracted by the robot and curious to see its reactions, how it would move, what it says, while very little interest was prompted by the tablet. We also offered participants the opportunity to try the vocal interaction with Pepper: although the users were very excited, it did not work very well in the end, probably due to the vocal features of some elderly users and also the difficulties that Pepper exhibited recognizing the Italian language spoken by people who (even occasionally) may have language-related difficulties. Therefore, in order to achieve fluent interaction, we decided not considering vocal input in this version of the application. During such initial tests, MCI patients showed a high level of interest and curiosity towards Pepper, and enthusiasm towards the possibility of interacting with a robot for their training. Figure 3 was taken during one of such tests: as you can see, the users' body postures reflect the general user interest towards Pepper.

Figure 3: Older adults with the humanoid robot Pepper

In the second phase, the users were divided into two different groups, balanced according to their computer experience by the psychologists of the clinics. One group of seven participants was assigned to the group interacting with the Pepper robot, while the remaining seven were assigned to the 'tablet' group. Our experimental test was included in the routine training session in the clinic. The game was played two days a week.

During this phase each individual participant interacted and played the music quiz game one by one. Before the start of the first test session, we asked the older adults to compile two pre-study questionnaires: A computer experience questionnaire, and a demographic questionnaire. In addition, after the first session, they had to fill in the User Engagement Scale questionnaire (O'Brien et al., 2018), which was also repeated after the last game session. After the compilation of the two questionnaires we asked them to log in, then listen to the instruction given by the device and then recognize the title or the singer of the 15 songs proposed by the game.

The users were called one by one in a dedicated room for their own session training class where they were observed by one moderator. The Pepper and Tablet sessions were done simultaneously in different rooms.

In the sessions, the users were at a distance of 30-40 cm from the robot, so that they were close enough to interact with it and the robot's animations with its arms did not bother the participants. The participants who interacted with the tablet were seated in a chair next to a desk where the tablet was located.

Participants worked without a maximum time limit. We measured: number of correct answers, wrong answers, number of answers not provided within the time limit, time needed to complete the whole game, the time needed to respond to each answer and the number of trials needed to provide 12 correct answers (which corresponds to 80% of the total number). For each group, a moderator was available during the various sessions, and took note of user feedback or any significant events occurring during user interactions with the robot and the tablet.

Results

Completion Time and Task Success

The data show that the completion times was slightly better for the robot (see Figure 4).

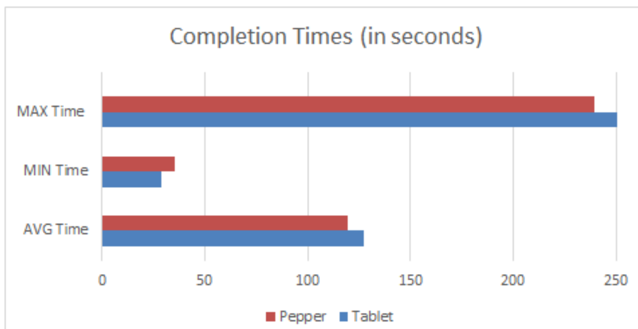

Figure 4: Completion Time in the two conditions

Figure 5 shows the evolution of the average completion times over sessions in the two conditions (Tablet vs. Pepper). On the X axis the sessions are visualised, whereas on the Y axis the time is reported in seconds. As it can be seen from such figure, under both conditions, the tendency was that the average completion time decreased over time, assuming quite stable values approximatively in the mid of the training.

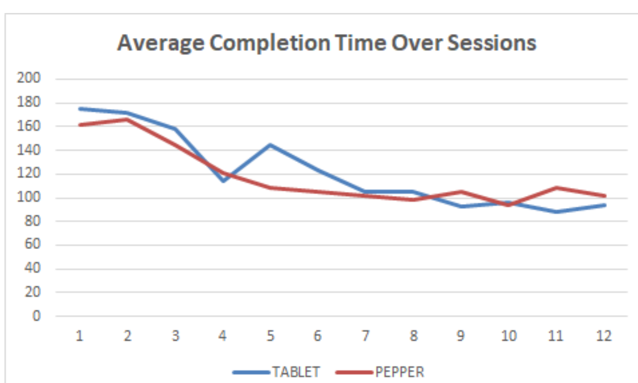

Figure 5: Average completion times over sessions in the two conditions (Tablet vs. Pepper)

The number of the correct answers provided during the interaction with the game was slightly better for the tablet. On average, in the Pepper group, each user correctly answered 10.6 questions, whereas in the Tablet group each user correctly answered 11.2 questions. In both versions, the minimum number of correctly answered questions by users was 4, the maximum number was 15.

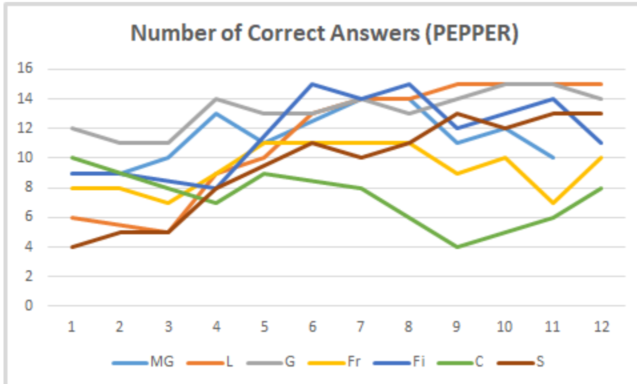

Figure 6: Number of Correct Answers Over Sessions (Pepper)

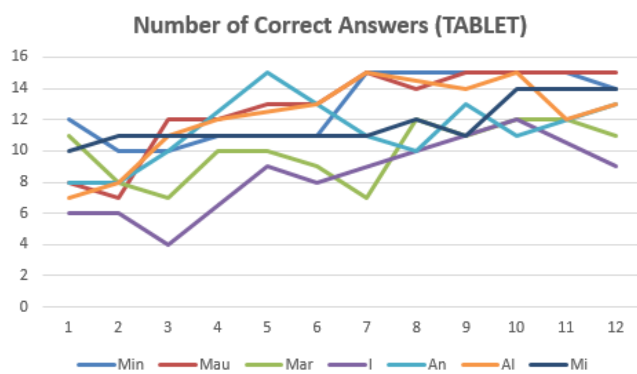

Figure 7: Number of Correct Answers Over Sessions (Tablet)

Figure 6 and Figure 7 show the evolution of the number of correct answers per user. As it can be seen from such figures, overall, under both conditions, the general tendency was that this number increased over time.

Regarding the answers that users were not able to provide by the time available, they were on average 0.7 answers on the tablet and 0.4 on Pepper. As for the answers that were provided but they were wrong, they were 3.1 on tablet and 4 on Pepper, on average.

We also consider the number of users who were able to correctly answer at least 80% of the provided questions (=12 questions). For each of such users, we calculated the number of trials needed to answer such 12 questions correctly. For the tablet version users on average needed 5 trials, whereas on the Pepper platform, users on average needed 4.4 trials, which is slightly better.

User Engagement

In the User Engagement Scale questionnaire, the subscales (1 strongly disagree, 5 completely agree) considered are: FA= Focused Attention (to what extent the application is able to receive focused attention by users); PU= Perceived Usability (the usability of the application as perceived by the user); AE= Aesthetics (to what extent the application is aesthetically attractive); RW= Reward (to what extent the user experience of interacting with the application is rewarding for the user).

Table 1 reports the user engagement values collected at the beginning of the test (i.e. after the end of the first session). As it can be seen, the subscale that received the highest value at the beginning of the test was Reward for the tablet and Perceived Usability for Pepper. At the beginning of the test, the overall, composite engagement score on tablet was calculated as 16.1, for Pepper it was 16.2.

Table 1: UES values collected at the beginning of the test (after the first session)

Subscale	M(SD) Tablet	M(SD) Pepper
FA	3.5 (0.8)	3.4 (1.1)
PU	4.1 (0.7)	4.5 (0.8)
AE	4.2 (1.0)	4.2 (0.5)
RW	4.3 (0.3)	4.1 (0.7)

Table 2 shows the user engagement values collected at the end of the test (i.e. after the end of the last session). As Table 2 shows, the subscale that received the highest value at the end of the test was still Reward for the tablet and (again) Perceived Usability for Pepper. At the end of the test the overall, composite engagement score was calculated as 16.2 for tablet, 16.7 for Pepper.

Table 2: UES values collected at the end of the test (after the last session)

Scale	M(SD) Tablet	M(SD) Pepper
FA	3.3 (0.5)	3.1 (0.8)
PU	4.2 (0.8)	4.6 (0.3)
AE	4.2 (0.7)	4.5 (0.5)
RW	4.5 (0.6)	4.5 (0.5)

Therefore, although no particular differences were revealed between the two conditions, participants rated their engagement as higher in the robot condition than in the tablet condition. Thus they were more engaged by the robot –based version of the application. Therefore, the more empathically/emotionally rich feedback provided by the robot could have affected the level of their engagement.

User's feedback

During the game session of each participant, we took note about how they felt while interacting with the devices. At the beginning, the participants that discovered not to be assigned to the 'only-tablet' group felt a bit disappointed for not having the possibility to interact with the robot. Nonetheless after experiencing the game, the caregivers reported that the game was well received by all participants.

Before starting the actual game, participants in both groups received an introduction in which the application greeted the user and then presented to the user the goal of the game. On the one hand, as soon as such introduction finished, the participants in the 'tablet' group immediately started the game session. On the other hand, users in the robot group responded to the robot's greeting spontaneously saying things such as "Good morning Pepper, you are so cute" and after the brief game introduction said by the device they typically answered "Thank you Pepper" and several of them touched the robot's hands. It is worth noting that this situation did not occur only in the first session, but it was repeatedly observed by the moderator over various sessions. In addition, some patients talked to the robot as it was one of their best friends, i.e. telling it information about their current physical/cognitive status, and/or expressing the willingness/pleasure to meet it. For instance, one participant said: "Hello Pepper, this morning when I woke up I had a severe back pain and I was about to decide not coming here, however, in the end I came because I knew I would have played with you". Furthermore, the patient said that the robot is very 'cute' and its behaviour encourages him to continue playing the game. So, in this case the robot was stimulating in making the older adults go out their house and more properly adhering to the training program, and it was approached in a way that resembles the kind of communication occurring among humans.

Even during the actual game, users of the robot group, differently from the others, continued considering the robot as having human traits: (e.g. they said that it has a cute face, also touching the hands and the arms of the robot). As another example, when a user was wrong in giving an answer, he told the robot:" I'm a bit slow, Pepper, please have patience with me". Occasionally, when they realised that the answer they chose had some similarities with the correct answer (for instance, it was a longer version of the correct title, or it was part of the refrain of the song but not the correct answer), they reacted as the robot itself would have tried to be misleading towards them ("You are trying to trick me"). On the other hand, one user, after correctly replying to an answer and after seeing the robot as expressing great enthusiasm, interpreted the robot's reaction as really empathic towards him ("Can you see? The robot is happy! If the robot is happy I'm happy too").

The game was well appreciated by all users, and resulted in a very competitive atmosphere. This was also perceived by caregivers who noticed that users sometimes discussed and also compared their game results with the others'. In addition, during the tests, some patients were observed as putting in place compensatory strategies for their memory lacks: for instance, one user tried to remember that the correct answer of a specific question was the one having the longest title between the possible options, instead of trying to memorise the correct answer. One patient of the 'tablet' group even brought to the training session a piece of paper where she had written down a few answers she declared unable to recall (of course, she was requested not to use it during the game). Caregivers reported that the robot significantly contributed to also make the atmosphere in the training program very dynamic, activating, and stimulating, even in a part of the training period (i.e. towards the end) when, traditionally, some drops in patients' attendance were registered during the previous editions. According to what caregivers reported, patients also well received the fact that, differently from previous editions of the TTB project where the games were basically group-based games, in this edition, it was a single-user game, therefore they could engage individually with it, and at their own pace. The inclusion of a highly innovative device like the robot, and the fact that each of them was asked to interact with it individually, increased their sense of 'importance' and self-esteem (which is generally low in this kind of patients). Finally, all the users successfully completed the training session, which was also reported as a good result per se.

DISCUSSION

Overall the test results were encouraging: users were enthusiastic about participating, and also caregivers expressed satisfaction about how this experimental training went and the level of engagement that users showed. As a result of this study, a number of design-related considerations can be identified.

Often it is difficult to motivate MCI users to train frequently and for a sufficiently long time using traditional training exercises which after a while are typically found repetitive. The provision of a digital game through a robot seems to successfully stimulate MCI patients in better committing to their training throughout its entire duration. In addition, the fact that users were positive toward Pepper and open to interacting with it, shows that older age is still compatible with robot use.

Furthermore, although not having a specific previous experience in interacting with robots, all the subjects were able to individually interact with Pepper, and successfully complete the training without experiencing particular difficulties. This shows that this type of device and training game can be considered as a viable and usable opportunity to be used with MCI elderly, also because it provides some emotional support, which is an aspect to which older adults are susceptible.

Another aspect we noted was the fact that the level of engagement in the group of people who interacted with the robot seemed to also affect the group that used the tablet. The game created quite a socially active and ‘competitive’ environment. While users of both groups were observed as interested to know and improve their own performance over time (when they received the final feedback at the end of a session they tried to compare it with the one they had got in the previous session, when they were able to remember it), especially at the beginning of the training the users of the robot group were observed as those more prompting discussions with other users about the obtained game results, to compare them with other users, which subsequently also affected users of the other group. Thus, the game also offered a good opportunity for socialization among the recruited patients.

Regarding the relationships between user engagement and performance in carrying out the submitted tasks, our initial hypothesis was that more engaged users would score higher in terms of correct answers. Instead, the results gathered in this evaluation show that more engaging systems do not necessarily foster better task success outcomes, as users of the (less-engaging) tablet version, on average, performed slightly better than users of the robot’s (see “Results” sub-section), in terms of correct answers provided.

A result of the study was that the more empathic feedback provided by the robot could have improved the level of the users’ engagement. In order to further improve the empathic aspects of the training and make it even more stimulating for the target subjects, further improvements can be envisaged. For instance, a more personalised interaction could be designed, starting from the feedback messages that the application provides to the user. Indeed, in our study we observed that people were very interested in their (and in the others’) performances, thus, more personalised feedback can be provided to keep users motivated and engaged. For instance, the feedback could include reinforcement phrases referring to previous user results or the average performance of the class, encouraging subjects to maintain and progress beyond their last played level and therefore stimulate users in the cognitive training.

Regarding task completion times, in the robot condition, Pepper adds some more empathic/emotional feedback (through body movements and lights) to better emphasize and reinforce the results. Therefore, taking into account the kind of target users involved in the test, it is worth pointing out that in the robot condition, users were able to stay focused on the tasks, even though they often needed to switch their attention between the two parts of the robot: the one providing graphical information (the chest) and the parts providing lively and emotional expressions (i.e. a mix of eye led colours, head postures, robot’s gestures and vocal feedback). Thus, we can argue that the emotional feedback provided by the robot did not distract them from their main task.

CONCLUSIONS AND FUTURE WORK

In the context of interventions for reducing cognitive decline in the elderly population, technologies have been increasingly conceived as a support for patients, their caregivers and the clinicians. The presented study aims to investigate how seniors with Mild Cognitive Impairment relate with and perceive serious games accessed through humanoid robots, as part of a training programme aimed to improve their cognitive status. One limitation of this work is that the interaction period was limited and the sample of subjects was not very large, therefore for future work we plan further evaluations over longer periods of time and considering larger trials, which can provide data suitable for statistical analysis as well. Another aspect that we plan to investigate in future work is the possibility to consider additional, more varied multimodal interactions to further increase the level of empathic support provided by the robot.

ACKNOWLEDGEMENTS

We thank all the volunteers that participated in the study, and help from anonymous for the implementation of the game.

REFERENCES

Bechade L., Dubuisson-Duplessis G., Pittaro G., Garcia M., Devillers L. 2019. Towards Metrics of Evaluation of Pepper Robot as a Social Companion for the Elderly. In: Eskenazi M., Devillers L., Mariani J. (eds). Advanced Social Interaction with Agents. Lecture Notes in Electrical Engineering, vol 510. Springer, Cham. https://doi.org/10.1007/978-3-319-92108-2_11

Chan M.Y., Haber S., Drew L.M., Park D.C. Training Older Adults to Use Tablet Computers: Does It Enhance Cognitive Function? The Gerontologist, Volume 56, Issue 3, June 2016, Pages 475-484. https://doi.org/10.1093/geront/gnu057

Churamani, N., Sutherland, A., Barros, P. 2018 An Affective Robot Companion for Assisting the Elderly in a Cognitive Game Scenario. In Proceedings of the Workshop on Intelligent Assistive Computing (WIAC), IEEE World Congress on Computational Intelligence (WCCI) 2018.

Cohen-Mansfield J., Dakheel-Ali M., Marx M.S. 2009. Engagement in persons with dementia: the concept and its measurement. Am J Geriatr Psychiatry. 2009;17(4):299–307. doi:10.1097/JGP.0b013e31818f3a52

De Almeida, R. X. E., Ferreira, S. B. L., and Soares, H.P. Recommendations for the Development of Web Interfaces on Tablets/iPads with Emphasis on Elderly Users, In Procedia Computer Science 2015, 140-149. https://www.sciencedirect.com/science/article/pii/S187705091503104X

Gauthier S., Reisberg B., Zaudig M., Petersen R.C., Ritchie K., Broich K., Belleville S., Brodaty H., Bennett D., Chertkow H., Cummings J.L., de Leon M., Feldman H., Ganguli M., Hampel H., Scheltens P., Tierney M.C., Whitehouse P., Winblad B. 2006. Mild cognitive impairment. In The Lancet. Vol 367:1262-70. doi: 10.1016/S0140-6736(06)68542-5.

Hoffman R., and Hancock, P., 2015. The Cambridge Handbook of Applied Perception Research, Cambridge: Cambridge University Press, doi:10.1017/CBO9780511973017.

Joddrell, P., Astell, A. J., Studies Involving People with Dementia and Touchscreen Technology: A Literature Review. JMIR Rehabil. Assist. Technol. 2016;3(2): e10. Published 2016 Nov 4. doi:10.2196/rehab.5788

Johnson, J., and Finn, K., 2017. Designing user interfaces for an aging population, Morgan Kaufman

Manera V, Petit P.D., Derreumaux A., Orvieto I., Romagnoli M., Lyttle G., David R., Robert P.H..2015. 'Kitchen and cooking,' a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. Frontiers in aging neuroscience vol. 7 24. 17 Mar. 2015. doi:10.3389/fnagi.2015.00024

McCallum S., Boletsis C. (2013) Dementia Games: A Literature Review of Dementia-Related Serious Games. In: Ma M., Oliveira M.F., Petersen S., Hauge J.B. (eds) Serious Games Development and Applications. SGDA 2013. Lecture Notes in Computer Science, vol 8101. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40790-1_2

Mentis, H.M., Madjaroff, G., and Massey, A.K. 2019. Upside and Downside Risk in Online Security for Older Adults with Mild Cognitive Impairment. In Proceedings of the 2019 Conference on Human Factors in Computing Systems (CHI '19). ACM, New York, NY, USA, Paper 343, 13 pages. https://doi.org/10.1145/3290605.3300573

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, Volume 112, 2018, Pages 28-39, ISSN 1071-5819. https://doi.org/10.1016/j.ijhcs.2018.01.004.

Park, K., Goh, T., and So, H.-J., Toward accessible mobile application design: developing mobile application accessibility guidelines for people with visual impairment, In Proceedings of HCI Korea 2014, 31-38. https://dl.acm.org/citation.cfm?id=2729491

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. In International Journal of Social Robotics (2019), pp 1–13. https://doi.org/10.1007/s12369-019-00533-y

Savulich, G., Piercy, T., Fox, C., Suckling, J., Rowe, J.B., O Brien, J.T., and Sahakian, B.J. 2017. Cognitive Training Using a Novel Memory Game on an iPad in Patients with Amnestic Mild Cognitive Impairment (aMCI). The international journal of neuropsychopharmacology vol. 20,8 (2017): 624-633. doi:10.1093/ijnp/pyx040

Schroeter, Ch., Mueller, S., Volkhardt, M., Einhorn, E., Huijnen, C., van den Heuvel, H., van Berlo, A., Bley, A., Gross, H.-M. 2013. Realization and user evaluation of a companion robot for people with mild cognitive impairments. In IEEE International Conference on Robotics and Automation (ICRA), 2013, pp. 1153-1159. doi: 10.1109/ICRA.2013.6630717

Stogl, D., Armbruster, A., Mende, M., Hein, B., Wang, X., and Meyer, P. 2019. Robot-Based Training for People with Mild Cognitive Impairment. In IEEE Robotics and Automation Letters, vol. 4, no. 2, pp. 1916-1923, April 2019. doi: 10.1109/LRA.2019.2898470

Tapus, A., Tapus, C., Mataric, M.J., 2009. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In IEEE International Conference on Rehabilitation Robotics (ICORR), Kyoto, 2009, pp. 924-929. doi: 10.1109/ICORR.2009.5209501

Vaportzis E., Martin M., Gow A.J. 2017. A Tablet for Healthy Ageing: The Effect of a Tablet Computer Training Intervention on Cognitive Abilities in Older Adults. In The American journal of geriatric psychiatry: official journal of the American Association for Geriatric Psychiatry vol. 25,8 (2017): 841-851. doi: 10.1016/j.jagp.2016.11.015

Wu, Y.H., Cristancho-Lacroix, V., Fassert, C., Faucounau, V., de Rotrou, J., Rigaud, A.S. 2016. The Attitudes and Perceptions of Older Adults with Mild Cognitive Impairment Toward an Assistive Robot. Journal of Applied Gerontology, vol 35, Issue1, 3-17. https://doi.org/10.1177/0733464813515092

W3C. Accessibility requirements for people with low vision. Retrieved September 06, 2019 from http://w3c.github.io/low-vision-a11y-tf/requirements