Feasibility Study on the Role of Personality, Emotion, and Engagement in Socially Assistive Robotics: A Cognitive Assessment Scenario

Alessandra Sorrentino 1,2, Gianmaria Mancioppi 1,2, Luigi Coviello 1,2, Filippo Cavallo 1,2,3, Laura Fiorini 3,*

Abstract: This study aims to investigate the role of several aspects that may influence human–robot interaction in assistive scenarios. Among all, we focused on semi-permanent qualities (i.e., personality and cognitive state) and temporal traits (i.e., emotion and engagement) of the user profile. To this end, we organized an experimental session with 11 elderly users who performed a cognitive assessment with the non-humanoid ASTRO robot. ASTRO robot administered the Mini Mental State Examination test in Wizard of Oz setup. Temporal and long-term qualities of each user profile were assessed by self-report questionnaires and by behavioral features extrapolated by the recorded videos. Results highlighted that the quality of the interaction did not depend on the cognitive state of the participants. On the contrary, the cognitive assessment with the robot significantly reduced the anxiety of the users, by enhancing the trust in the robotic entity. It suggests that the personality and the affect traits of the interacting user have a fundamental influence on the quality of the interaction, also in the socially assistive context.

Keywords: social assistive robotics; cognitive assessment; quality of interaction

1. Introduction
The current trend of the Human–Robot Interaction (HRI) field is to endow robotic platforms with social intelligence to create a more trustworthy and humanlike interaction, for example, by mimicking the human’s behavior [1]. One step towards the development of such capability is to integrate into the robotic platform a complete understanding of the profile of the person it is interacting with [2], which allows the robot to adapt its behavior based on the user preferences and needs. Advances in artificial intelligence, and machine learning (ML) mostly, aim at endowing robotic systems with powerful capabilities for interacting [3]. For instance, the work in [4] presents a personalized ML model with the aim of automatically detect the affective states and engagement of children involved in robot-assisted therapy. As reported in [3,5], computation personalization of HRI based on deep learning is still an open challenge due to the real-world complexities and the no clearness of the aspects the robot should be able to perceive.

Among the fundamental aspects that the robot should take into account, the personality of the user should be included, since it has been considered as a vital factor in understanding the quality of HRI [6]. As described in [7], personality represents a coherent pattern of behaviors that could be used to explain how people interact with others in social settings. Similarly, how a person behaves during the interaction could be affected by temporal mental states which are reflected by emotions and different levels of engagement. Both temporal traits (i.e., emotion and engagement) may fluctuate during the interaction...
or change their temporal dynamics based on some salient events that may happen during the interaction [8]. In HRI, several works [9–11] focused on inferring the engagement of a person by analyzing the non-verbal cues (e.g., nodding, mutual and direct gaze, the timing of the response to stimuli) of the human beings interacting with the social robots. When the interaction occurs with the elderly population, additional aspects of the profile may be considered, such as possible cognitive and physical impairments.

Given the advances in the communication modality of recent robotic platforms, more often social robots are used in assistive contexts as support for the assessment and the cognitive stimulation of elderly people [12–18]. Robots can be programmed to perform specific and repetitive actions with the advantage of performing them following the required standardization [16]. Compared to technologies usually used in this context, such as tablets and Personal Computers [19], socially assistive robots result more empathetic and challenging. This approach provides support to professional caregivers, allowing multiple screening and alleviating their workload while promoting sociality among the older population. Unlike previous works for cognitive assessment with robot technology [15,16], this study exploits the usage of a non-humanlike robot, ASTRO. It is endowed with simple social behaviors, such as speech synthesis (verbal behavior) and led illumination (non-verbal behavior).

Similar to [20], we would like to explore the role of the main factors which could affect the interaction of ASTRO with elderly people: Relatively permanent qualities (i.e., cognitive profile and personality) and temporal traits (i.e., emotional and engagements affects) of the interacting elderly users. Particularly, this feasibility study aims to investigate the role of each factor (i.e., cognitive profile, personality, emotional and engagement state) in the perception of the robot, its usability, and in the interaction itself. Our research questions could be summarized as follows:

1. Do the attitude of the users and the user’s perception of the technology change after the interaction with the robot? (RQ1).
2. Does the cognitive mental state influence the usability and the user’s perception of the robot in this scenario? (RQ2).
3. Do the personality traits influence the usability and user’s perception of the robot in this scenario? If yes, which one? (RQ3).
4. Does the current emotion influence the usability and user’s perception of the robot in this scenario? If yes, which one? (RQ4).
5. Which factors may influence the interaction in the assistive scenario? (RQ5).

Finally, this work discusses the results suggesting some guidelines for developing social assistive robots able to adapt and personalize their behavior based on the interacting participant in an assistive context.

2. Related Works

Socially assistive robotics (SAR) aims to assist human users by creating effective and close interactions to achieve measurable progress in convalescence and rehabilitation [21]. In healthcare intervention, robotic solutions have been designed as a mechanism to support mental health activities ranging from screening to administrating on-demand therapy [22]. In the latest years, several solutions adopted a robotic platform as support of cognitive therapy, assessment, and stimulation [23]. As reported in Table 1, the robotic system is usually programmed for administrating the overall test to older patients, under the supervision of a professional clinician. Human-like robots are mostly used (i.e., NAO, Pepper). Due to their appearance, human-like robots are thought to facilitate social interaction and communication, as they possess all the necessary features to convey social signals [17]. As shown in Table 1, the authors in [24] adopted a no-humanlike robot (i.e., Giraff robot) for cognitive stimulation. Namely, the robot was not endowed with special social behaviors and it was not able to give a feedback on the elderly performance [24]. Differently from [24], the robot adopted in this study resembles the human body shape (i.e., heads with eyes, torso). Furthermore, it is endowed with some communicative skills which allow it to motivate the user and to exhibit some social behaviors, as will be described in Section 3.2.
| Reference | Year | Robot | Aim | Robot Role | Human Role | Outcome |
|-----------|------|-------|-----|------------|------------|---------|
| [13]      | 2009 | Bandit | Cognitive Therapy | Guiding the patient in performing the game | Supervision | Robot encouragement improves response time. |
| [14]      | 2017 | NAO   | Cognitive Therapy | Guiding the patient in performing the game | Supervision | Robot acceptability increases after the interaction. |
| [15]      | 2018 | Pepper | Cognitive Assessment | Overall administration of the test | Supervision | Robot improved socialization. |
| [16]      | 2019 | Pepper | Cognitive Assessment | Overall administration of the test (1st time) | Overall administration of the test (2nd time) | Validation of the robotic assessment. |
| [17]      | 2019 | NAO   | Cognitive Therapy | Overall administration of the test (1st time) | Overall administration of the test (2nd time) | Analysis of nonverbal behavior revealed there was more engagement with the robot, than with the clinician. |
| [24]      | 2019 | Giraff | Cognitive Stimulation | Overall administration of the test | Supervision | Older people accepted the guidance of a robot, feeling comfortable with it explaining and supervising the tests instead of a clinician. |
| [25]      | 2020 | NAO   | Cognitive Stimulation | Overall administration of the test (1st time) | Overall administration of the test (2nd time) | Performing stimulation exercises with the robot enhanced the therapeutic effect of the exercise itself, reducing depression-related symptoms in some cases. |
| [26]      | 2021 | Pepper | Cognitive Therapy | Guiding the patient in performing the game | Supervision | Robot improved socialization. |
| This work | 2021 | ASTRO | Cognitive Assessment | Overall administration of the test | Supervision | Robot reduced anxiety and incentivized the interaction. |

The main success of the adoption of robotic solutions in this context relies on the enhanced engagement of the participants concerning the non-embodied computer screen [16]. In previous works in the SAR field, the engagement is usually evaluated through questionnaires (i.e., User Engagement Scale questionnaire in [26]), non-behavioral cues (i.e., mainly eye gaze direction [9–11,25]), and physiological signals (i.e., EEG and galvanic skin response [14]). Similarly, the authors of [17] evaluated the emotional experiences of the users undergoing a cognitive stimulation activity with NAO. They investigated the two dimensions of emotional experience: Valence and arousal.

While the first one was assessed using the self-report Positive and Negative Affect Schedule (PANAS) questionnaire, the latter dimension was evaluated by analyzing the heart rate variability. Additionally, the authors evaluated offline the non-verbal behaviors (i.e., face-directed gaze, gaze aversion, hand movement, and posture) of the participants to assess their engagement level. The authors of [25] investigated the emotional experience characterizing the interaction with the robot, by deriving the positive emotions from the number of times and the duration of smiling activity occurring during the experimentation. The same authors investigated also the usage of Facial Action Units (FACs) for detecting happiness, surprise, anger, fear, disgust, sadness, and neutrality in more recent work.
in the same context [18]. This brief regression on the temporal traits (i.e., emotion and engagement) analyzed in previous works enforces our idea that they should be included as factors that may influence the quality of the interaction. In this work, we will investigate the older adults’ emotional and engagement response in terms of valence (i.e., by adopting the PANAS questionnaire) and by analyzing the duration and the frequency of non-verbal behaviors.

As described in [27], emotions could be generated by personal affective disposition and relatively permanent affective personality traits. It is the reason behind the current trend of HRI in investigating the influence of personality on the quality of the interaction [6]. Indeed, personality traits could be used to explain the way people respond to others in social settings. Among the several models of personality, the Big Five personality traits [28] are the most widely used. They include extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience. As highlighted by the review reported in [6], traits of the personality influence the user’s behavior during the interaction. In the SAR field, the work described in [15] represents the first attempt of investigating the influence of the user’s personality traits on psychometric assessments. Also, in this case, the traits of the Big Five model were used and the openness to experience aspect was the one that influenced the final cognitive score. Similarly, our work aims to investigate the role of the personality factors, focusing on their influence on the quality of the interaction more than on their effect at a cognitive level. Our ground hypothesis is that both personality and cognitive mental state may influence the quality of interaction. They are both aspects that the robots should take into account to modulate the interaction appropriately.

In a similar fashion to the works reported in Table 2, this work will also analyze the usability and the perception of the technology, by collecting the feedback directly from the interacting participants. A series of self-report measures will be included to evaluate the state of anxiety perceived by the user, as well as to identify the perception of this technology by non-expert users.

### Table 2. List of investigated factors in previous socially assistive robotics (SAR) works.

| Reference | Year | Investigated Factors          | Tools                              | Feedback                                      |
|-----------|------|--------------------------------|-----------------------------------|-----------------------------------------------|
| [13]      | 2009 | -                              | -                                 | A short questionnaire on the experience       |
| [14]      | 2017 | Engagement                     | Physiological signals (EEG, GSP)  | Robot User Acceptance Scale-RUAS              |
| [15]      | 2018 | Personality                    | NEO Personality Inventory-3       | Unified Theory of Acceptance and Use of Technology (UTAUT) |
| [17]      | 2019 | Engagement and Emotion         | PANAS, video analysis, and        | NASA Task Load Index                          |
|           |      |                                | Physiological signals             |                                               |
| [24]      | 2019 | -                              | -                                 | Usability questionnaire                       |
| [25]      | 2020 | Engagement and Emotion         | Video analysis                    | State-Trait Anxiety Inventory, Psychosocial Impact of Assistive Devices Scales, and System Usability Scale |
| [26]      | 2021 | Engagement                     | User Engagement Scale questionnaire, video analysis | -                                             |
| This work | 2021 | Engagement, Emotions, Personality, Cognitive Mental state | PANAS, video analysis, BFI, and MMSE | Godspeed questionnaire, Ad-hoc usability questionnaire |

### 3. Materials and Methods

#### 3.1. Cognitive Assessment

In geriatric health care, there are several screening tools for cognitive impairment, such as the Mini-Mental State Examination (MMSE) [29] and the Montreal Cognitive Assessment (MoCA) [30]. In this work, we focused on the former one, since it was the one adopted by the professional caregivers collaborating with the study. The MMSE encompasses 21 items
which cover tests of recall, naming, registration, orientation, writing, reading, drawing, comprehension, repetition, calculation, and attention [29]. The cognitive status level is obtained by summing the scores of each category and normalizing it based on the age and educational level of the participant. In detail, the cognitive status level could fall into one of the following categories: Severe cognitive impairment, mild cognitive impairment, and no cognitive impairment.

3.2. The Robot

The robot adopted in this study was the ASTRO robot [31]. It is a robot that supports personal mobility, which supports the Robot Operating System (ROS) as middleware. It is mounted on a SCITOS G5 robotic platform (Metralabs GmbH, Germany) and it is equipped with several sensors (i.e., laser, RGB-D cameras, microphones, speakers, and force sensors) and two tablets, one on the front and one on the back. As shown in Figure 1, it is a non-humanoid robot, with limited gestures and tilting modalities. The only anthropomorphic aspect is given by the shape of the head, which incorporates two blinking Light Emitting Diodes (LEDs) inside.

Figure 1. The architecture of the ASTRO robot.

In the present work, ASTRO was used in a Wizard-of-Oz modality by a professional psychologist. To simplify the administration of the MMSE, we implemented a web interface that allowed the caregiver to guide the robot through the administration. Thanks to the rosbridge_server installed on the platform, the commands were directly forwarded to the robot controller from the interface. The controller was then responsible to publish the data over the dedicated topics to activate the speech and the blinking LEDs over the platform. To avoid any networked issue, the text-to-speech (TTS) implemented over the robotic platform was based on voice synthesizer Pico TTS. With this approach, the caregiver could make the robot repeat the questions and motivate or comfort the patients with ad-hoc sentences. On the contrary, the color blinking LEDs were controlled by a specific ROS node to manage the transparency of speaking turns during the interaction. Specifically, the green color was active when the robot was speaking, and the blue color was activated in listening mode. Furthermore, the web interface allowed the caregiver to collect the answers provided by the patient. In the case of the writing and drawing tasks, the robot asked the participant to show the results to the camera.
3.3. Experimental Procedure

The experimental procedure adopted in this study is composed of 3 stages, as shown in Figure 2b. In the first stage (T1), the participant was informed about the procedure and the functionality of the ASTRO robot. Then, the participant was asked to answer four questionnaires: One about his/her personality (i.e., Big Five Inventory), one about the current emotional state (i.e., PANAS), and the last two related the first impression about the robotic platform (i.e., GodSpeed Questionnaire) and its perceived usability (i.e., ad-hoc usability questionnaire), respectively. They will be described in the following section. During the second phase (T2), the robot introduced itself and welcomed the participant. To engage the user in the interaction, the robot administered the socio-demographic questionnaire at the beginning of T2. After it, the robot administered the MMSE (see Figure 2a) following the instructions provided by the psychologist and thanking the participant at the end. The overall administration was guided by the clinician through the dedicated web interface (see Section 3.2). To collect the feedback on the performed experience, during the final phase (T3), the operator administered three questionnaires: One about the emotional state of the user, one about the usefulness of the service, and one about the perceived capabilities of the robot. During the whole experimentation, a technician assisted the trial to be ready to intervene in case of necessity. The overall interaction between human–robot was recorded by the robot’s camera to analyze it offline.

![Figure 2. Experimental setup: (a) Elderly participant performing the drawing task of MMSE with ASTRO robot; (b) scheme of the experimental setup (T1–T3 are conducted by the professional therapist; T2 administrated by ASTRO robot in WoZ).](image)

3.4. Questionnaires

To investigate the personality of the participants, the Big Five Inventory (BFI) [32] has been used in this study. This test estimates five traits of the user’s personality: Extroversion (BFI_EXT), agreeableness (BFI_AGR), conscientiousness (BFI_CON), neuroticism (BFI_NEU), and openness to experience (BFI_OP). To reduce the overload of the target participant to complete the questionnaire, the short form of the BFI (i.e., BFI-10 [33]), which is composed of 10 items, was administrated in this study. On the other hand, the temporary emotional state of the user was assessed by using the Positive and Negative Affect Schedule (PANAS) [34] in the Italian version [35]. The questionnaire is composed of 20 adjectives that reflect positive attitudes (e.g., enthusiastic, determinate, curious) and negative ones (e.g., angry, hostile, nervous). This questionnaire was administrated in the “at this moment” timeframe at the beginning (P1) and at the end (P2) of each experimental session to investigate possible changes of humor related to the interaction with the robot. For both BFI-10 and PANAS questionnaires, the participants were asked to rate the characteristics that may or may not be applied to themselves on a 5-Likert scale (1 = strongly disagree, 5 = strongly agree).
The Godspeed questionnaire [36] was administrated to assess the perceived capabilities of the robotic platform. It is composed of 23 adjectives describing five key concepts on robot perception: Anthropomorphism (ANT), animacy (ANI), likeability (LIK), perceived intelligence (PEI), and perceived safety (PES). The participants were asked to rate the perceived capabilities on a 5-Likert scale. Furthermore, to investigate changes in the perception of the robot capabilities after the interaction, the Godspeed questionnaire was administrated before (GQ1) and after the experimentation (GQ2). To assess the usability of the robot and its service, the ad-hoc usability/acceptability questionnaire adapted by Cavallo et al. [37] was administrated. It evaluates the services provided by the robot and the positive or negative influence on the user’s mood (i.e., feeling of enjoyment or anxiety). It consists of a total of 9 items covering the following macro-aspects: Disposition about the services provided by the robot (ITU), feelings of anxiety (ANX), enjoyment (ENJ), and trust (TRU) evoked using the robotic platform, and perceived ease of use of the service (PEU). With respect to the original version described in [37], the modified version of the ad-hoc usability questionnaire was administrated at the beginning (US1) and the end (US2) of the interaction.

The general information of the users’ background was assessed utilizing the socio-demographic questionnaire, administrated by the robot in T2.

3.5. Participants

The experimental sessions took place in a residential facility in Tuscany located in Pistoia. The professional caregivers participating in the experimentation were responsible for the recruitment of the participants, among the ones who lived in the residential facility. The inclusion criteria were: Elderly users who did not have hard hearing and/or speaking impairments and who had an MMSE score > 10. A total of 11 participants were recruited for this feasibility study. All procedures were approved by the joint Ethical Committee of Scuola Superiore Sant’Anna and Scuola Normale Superiore (Delibera 14/2020). All the participants signed the consent form before entering the test.

3.6. Data Analysis

Data collected during the experimentation scenario were offline analyzed. As a preliminary step, the MMSE scoring was attributed by the professional psychologist who watched the recorded videos. While watching the recorded videos, the professional psychologist annotated behavioral parameters that could be used as descriptors of the engagement of the user during the interaction. As described in [38], the following categories of behavioral cues were considered:

- Verbal interaction quality metrics: Robot’s questions, robot’s quote, robot’s repetition, robot’s interaction time, interaction with the therapist, therapist interaction time, other time, total interaction time, coherent interaction, incoherent interaction;
- Emotion: Joy, neutral;
- Gaze: Direct, none;
- Facial expressions: Smile, laugh, frown, raised eyebrows, inexpressive; and
- Body gestures: Lifting shoulders, nodding head, shaking the head, quiet.

For each element of the category, the time spent and its percentage, based on the complement activities, were computed for each user. A simple descriptive statistic was calculated for each subcategory to evaluate the trend and a Spearman correlation analysis was performed to detect relations with the questionnaires’ responses.

The analysis of the questionnaires was then conducted. The total score of the PANAS questionnaire was computed by summing up the answers to all the questions. Furthermore, the score of the Positive and Negative Affects (PA, NA respectively) was computed by adding up only the items belonging to the corresponding category, respectively. Descriptive statistics were used to evaluate the PANAS scores and the domains of the Godspeed and ad-hoc usability questionnaires. Due to the small dimensions of the sample non-parametric statistic was used. The Mann–Whitney U test was applied on the two versions of the
PANAS, Godspeed, and Usability questionnaires to check if there were differences in the answers before and after the interaction with the robot (RQ1). Spearman correlation index ($\rho$) was computed to find any correlation between MMSE score and the Godspeeds’ questionnaires and/or the usability’s domain (RQ2). Similarly, the influences of the BFI personality traits and PANAS scale on the usability and perception of the robot were assessed with the Spearman correlation index (RQ3 and RQ4). To answer RQ5, we performed the correlation of the extracted engagement parameters with the BFI, MMSE, and PANAS scores, respectively. The assessment of the size of the correlation coefficients is based on the threshold values described in [39].

4. Results

Among the recruited participants, two users were excluded from the study due to a crying lag and an extremely delusional state in T1, respectively. As result, a total of 9 elderly persons (3 male, 6 female, avg. age 83.33 years old, range 72–92 years old) were enrolled for this study. Unfortunately, one participant refused to answer the questionnaires administrated in T1 and T2, while accepting to interact with the robot; another participant was not recorded by the camera, due to a technical issue. Based on the difficulty encountered, each user was excluded by the analysis involving the missed information. The Big Five personality traits of the recruited participants are summarized in Figure 3.

![Big Five personality traits of the recruited participants.](image)

By taking into account three levels of cognitive impairments (i.e., none, mild, severe), the pool of participants was composed of two users with severe cognitive impairments, two participants with no cognitive impairments, and the remaining participants with mild cognitive impairments.

4.1. Video Recording Analysis

On average, the time spent in each session was 11.21 min (std = 2.02 min). The average percentage of time spent interacting with the robot corresponded to 82.9% (std = 9.9%), which is more than the percentage of time spent interacting with the therapist (avg: 16.6%, std: 9.5%). Indeed, the average number of times the participants interacted with the therapist was 7.5 (std: 4.2). Mostly the interactions were characterized by coherent responses of the elderly (avg: 23.9, std: 7.3), and very few robot repetitions were recorded (avg: 5.7, std: 3.5). The analysis of the emotional attitude through no-verbal behaviors highlighted a general neutral emotional state (avg: 98.3%, std: 1.8%) characterized by a high percentage of inexpressiveness (83.6%). Despite the high percentage of inexpressiveness, the remaining facial expressions were also present at different extents: Smile (1.1%), laugh (0.5%), raised eyebrows (2.3%), and frown (12.5%). The same results were also confirmed by the body
posture analysis. Overall, 97.6% of the interaction was conducted in a quiet position, on average, and in rare cases, the people nodded their head in a sense of approval (1.4%) or shook their head in terms of disapproval (0.8%). Additionally, the gaze analysis reported that 42.7% of the time the users looked directly at the robot’s head during the interactions. Table 3 shows the detailed results.

Table 3. Descriptive statistics of the verbal and no-verbal parameters extracted from the video analysis.

| Domain                        | Acronym               | Measure   | Average | Standard Deviation |
|-------------------------------|-----------------------|-----------|---------|--------------------|
| Verbal interaction quality metrics | Robot’s question Count | 18.5      | 2.3     |
|                               | Robot’s Quote Count   | 15.5      | 1.7     |
|                               | Robot’s Repetition Count | 5.7   | 3.5     |
|                               | Coherent Interaction Count | 23.9 | 7.3     |
|                               | Incoherent Interaction Count | 2.3   | 3.6     |
|                               | Robot Interaction Time Time (sec) | 560 § | 146.5 |
|                               | Interaction with the therapist Count | 7.5  | 4.2     |
|                               | Therapist Interaction Time Time (sec) | 109.1 | 64.6 |
|                               | Other Time Time (sec) | 3.1       | 7       |
|                               | Total Interaction Time Time (sec) | 672.3 * | 121.2 |
| Emotions                      | Joy Time (sec)        | 12.6      | 1.7     |
|                               | Neutral Time (sec)    | 659.6     | 98.3    |
| Gaze                          | Direct Time (sec)     | 278       | 142.6   |
|                               | None Time (sec)       | 394.3     | 205     |
| Facial Expression             | Smile Time (sec)      | 8.1       | 1.1     |
|                               | Laugh Time (sec)      | 3.5       | 0.5     |
|                               | Frown Time (sec)      | 89.3 *    | 87.3    |
|                               | Raised Eyebrows Time (sec) | 14.1 | 16.7   |
|                               | Inexpressive Time (sec) | 557.3 | 105.3 |
| Body Gestures                 | Lifting Shoulders Time (sec) | 1.3   | 3.5     |
|                               | Nodding Head Time (sec) | 8.8     | 16.6    |
|                               | Shaking Head Time (sec) | 5.4 ◦   | 5.8     |
|                               | Quiet Time (sec)      | 656.9     | 127.5   |

* Significant differences between participants with different personality traits. ◦ Significant differences between participants with different emotional states. § Significant differences with the GQ2 scores. * Significant differences with the US2 scores.
The correlation analysis registered a very high positive correlation between the robot interaction time and the GQ2_ANI domain ($\rho = 0.93$). Similarly, a very high positive correlation was detected between Nodding head behavior and the US2_ITU ($\rho = 0.92$), suggesting that the quantity of time and some non-verbal behaviors influence the perception of the robot and its usability. In the following subsections, we will investigate the role of the engagements’ parameters with the users’ personality, rated affect, and cognitive mental profile.

4.2. Descriptive Statistics

Tendentially, the P1 results highlight a high PA (avg: 31.33, std: 6.94) and a low NA (avg: 14.16, std: 3.6). Similarly, the answers to P2 are characterized by a high PA (avg: 37, std: 3.6) and a low NA (avg: 13.16, std: 4.02). The results underline a small increment of the PA values and a small decrement of the NA scores after the interaction with the robot (P2). It results in a general trend of increasing the PANAS average score (avg P1 = 45.5, avg P2 = 50.16), which is not statistically meaningful ($p > 0.05$).

Similarly, the comparison of the GQ1 and GQ2 answers does not show any significant differences, even if there is an evident increasing trend for each item and macrodomains. Namely, the average score of the GQ1 is 61.57 points (std: 9.55), while the average score of the GQ2 is 71.71 points (std: 12.55). The descriptive statistic of each item is reported in Table 4.

Table 4. The average and the standard deviation of each item of the Godspeed questionnaires, administrated before (GQ1) and after (GQ2) the cognitive assessment.

| Domain     | Acronym | Item          | GQ1       | GQ2       |
|------------|---------|---------------|-----------|-----------|
| Anthropomorphism | ANT     | Natural       | 1.86 (1.46) | 2.71 (1.7) |
|             |         | Human-like    | 1 (0)     | 1 (0)     |
|             |         | Consciousness | 2.14 (1.67) | 2.57 (1.51) |
|             |         | Lifelike      | 1.28 (0.49) | 1.14 (0.37) |
|             |         | Moving        | 2 (1)     | 2.57 (0.78) |
|             |         | Elegantly     |           |           |
| Animacy    | ANI     | Alive         | 1.71 (0.75) | 2 (1.53)  |
|            |         | Lively        | 1.86 (1.07) | 1.42 (0.53) |
|            |         | Organic       | 1 (0)     | 1 (0)     |
|            |         | Interactive   | 1.57 (0.79) | 2.14 (1.34) |
|            |         | Responsive    | 1.86 (1.46) | 2.71 (1.89) |
| Likeability| LIK     | Like          | 4 (1)     | 3.85 (1.21) |
|            |         | Friendly      | 3.57 (1.4) * | 4.42 (0.78) |
|            |         | Kind          | 4.14 (0.9)  | 4.85 (0.37) |
|            |         | Pleasant      | 3.71 (1.11) | 4.14 (0.69) |
|            |         | Nice          | 3 (1)     | 3.57 (1.39) |
| Perceived  | PEI     | Competent     | 3.43 (1.27) | 3.71 (1.7)  |
| Intelligence|        | Knowledge     | 3.85 (0.69) | 4.28 (1.11) |
|            |         | Responsible   | 3 (1.63)  | 3.71 (0.75) * |
|            |         | Intelligent   | 3.42 (1.51) | 4 (1)     |
|            |         | Sensible      | 2.14 (0.89) | 3.42 (1.71) |
| Perceived  | PES     | Relaxed       | 3.71 (1.25) | 4.57 (0.78) |
| Safety     |         | Calm          | 3.71 (1.38) | 4 (1.73)  |
|            |         | Surprise      | 3.57 (1.51) * | 3.85 (1.67) * |

* Significant differences between participants with different personality traits. ◯ Significant differences between participants with different emotional states. § Significant differences with the MMSE scores.

Overall, the participants had a good disposition about the robot before and after the interaction (US1_ITU, avg: 10, std: 2; US2_ITU, avg: 10.8, std: 4.9). The perception of the easiness of usage of the platform changed from lower (US1_PEU, avg: 3.2, std: 0.45) to higher values of agreement (US2_PEU, avg: 4.4, std: 0.9). A less positive trend is registered.
on the feeling of enjoyment, which is characterized by middle values (US1_ENJ, avg: 3.2, std: 1.48; US2_ENJ, avg: 3.8, std: 1.6). In general, most of the participants did not feel any anxiety during the interaction with ASTRO (US1_ANX, avg: 5, std: 1.2; US2_ANX, avg: 2.2, std: 0.4) and trusted the robotic service (US2_TRU, avg: 6.8, std: 1.64; US2_TRU, avg: 7.8, std: 2.2). Among all, the Mann–Whitney U Test highlights a significant difference in the ANX domain before and after the interaction with the robot ($p = 0.02$). The descriptive statistic of each category is reported in Table 5.

Table 5. The average and standard deviation of each domain and item of the ad-hoc Usability questionnaires administrated before (US1) and after (US2) the cognitive assessment.

| Domain       | Acronym | Item                                                                 | US1       | US2       |
|--------------|---------|-----------------------------------------------------------------------|-----------|-----------|
| Disposition  | ITU     | “I would use the robot in case of need (i.e., if I was sick)” (Q1)    | 10 (2)    | 10.8 (4.9)|
|              |         | “I would be willing to use the cognitive service if it could help      | 4 (0.7)   | 4 (1.22)  |
|              |         | the family/caregiver’s work” (Q2)                                      |           |           |
|              |         | “I think my independence would be improved by the use of the robot”   | 2.8 (1.3) | 3.4 (1.81) *|
|              |         | (Q3)                                                                  | 3.2 (1.48) * | 3.4 (2.19) °|
| Anxiety      | ANX     | “I am too embarrassed in using the robot, around the community or the | 5 (1.22)  | 2.2 (0.4) |
|              |         | family” (Q4)                                                           | 3 (1.58)  | 1.2 (0.44) |
|              |         | “I am/was nervous doing the cognitive assessment with the robot”      | 2 (1)     | 1 (0)     |
| Enjoyment    | ENJ     | “I will enjoy/enjoyed using the robot for doing the cognitive         | 3.2 (1.48)| 3.8 (1.6) |
|              |         | assessment” (Q6)                                                       |           |           |
| Trust        | TRUST   | “I would trust in robot’s ability to perform the cognitive assessment”| 6.8 (1.64)| 7.8 (2.2) |
|              |         | (Q7)                                                                  | 3.6 (1.14)| 4 (1.41)  °|
|              |         | “I think the robot would be too intrusive for my privacy” (Q8)         | 3.2 (1.09)| 3.8 (1.64) |
| Perceived    | PEU     | “I found the robot easy to use to perform the cognitive assessment”   | 3.2 (0.45)| 4.4 (0.9) |
| Easy of      |         | (Q9)                                                                  |           |           |
| Use          |         |                                                                       |           |           |

* Significant differences between participants with different personality traits. ° Significant differences between participants with different emotional states.

4.3. Correlation with Personality

To investigate the role of personality with the perception of the robot, we computed the Spearman correlation with the items of GQ1 and GQ2. The correlation of the BFI aspects and GQ1 returned a negative correlation between GQ1_Surprise and BFI_AGR ($\rho = -0.72$), and the GQ1_Friendly item and BFI_NEU ($\rho = -0.74$). Negative correlations were also found between GQ2_Responsible and BFI_NEU ($\rho = -0.88$) and BFI.OP ($\rho = -0.93$). The BFI_OP was also negatively correlated with GQ2_Surprise ($\rho = -0.87$).

High negative correlations resulted by comparing BFI personality traits with the US1 and US2. In detail, the BFI_OP was negatively correlated with the ITU_Q3 (“I think my independence could improve with the robot”) of US1 ($\rho = -0.88$) and with the ITU_Q2 (“I would be willing to use the cognitive service if it could help the family/caregiver’s work”) of US2 ($\rho = -0.78$). These results show that the creative and inventive participants tend to disclaim the usage of the robot for improving their independence.

The personality traits of the older participants resulted correlated also with behavioral parameters extracted offline. As result, the total time of interaction was highly correlated with BFI.Ext ($\rho = 0.77$) and highly negatively correlated with the BFI_NEU ($\rho = -0.8$). It suggested that the more you are sociable and talkative, the more the experimental session lasted, and vice-versa. Surprisingly, the BFI_NEU trait was also negatively correlated with the frown facial expression ($\rho = -0.83$).
4.4. Correlation with Emotional State

The Spearman analysis reported one high correlation between the positive affection and GS1_Conscious item ($\rho = 0.76$) before the interaction with the robot. Similarly, the GS2_Conscious of the robot is very highly positively correlated with the negative affects of the participants after the interaction ($\rho = 0.95$). It suggests that the attribution of consciousness to a robotic platform can vary with the emotional states of the person interacting with. Furthermore, the GS2_Sensible is highly positively correlated with the total score of P2 ($\rho = 0.88$), suggesting that the more the person is sensible, the more he/she attributes the same condition to the robot.

The analysis between the emotional state of the participants and the perceived usability of the robot before the interaction did not report any significant correlation. On the contrary, after the interaction, the emotional state of the participants resulted very positively correlated with the disposition about the services (ITU) and the feeling of trust (TRU). Namely, the total score of P2 was highly correlated with US2_ITU_Q3 ($\rho = 0.92$) and US2_TRU_Q7 (“I would trust in robot’s ability to perform the cognitive assessment”) ($\rho = 0.97$). Very high positive correlations were registered combining the positive affect and the scores of US2_ITU_Q3 and US2_TRU_Q7. It suggests that people experiencing a “high degree” of emotions (i.e., mostly positive ones) tend to trust the interacting robotic platform and its usage.

This result is also confirmed by the correlations of emotional state with the behavioral parameters. Specifically, a very high negative correlation coefficient linked the total score of the P2 with the shaking head activity ($\rho = -0.97$). It can be interpreted as the more the person was sensible, less frequently expressed disagreement or felt upset at the interaction.

4.5. Correlation with Cognitive State

The MMSE score of each participant was used to investigate its influence on the perception of the robot, its usability, and on the engagement during the interaction. Regarding the perception of the robot, the results highlighted a very high negative correlation coefficient between the GS2_Conscious item and the cognitive state of the participants ($\rho = -0.98$). It suggests that participants with impaired cognitive capabilities tend to attribute a conscious capability to the social robot. It only occurs after the interaction with the robot, since the MMSE score does not significantly correlate with any item of the GS1. Similarly, the cognitive state of the participants did not influence either the perceived usability either the engagement behaviors during the interaction.

5. Discussion

This work presents a feasibility study that aims to investigate how persistent (i.e., personality) and temporal traits (i.e., emotion and engagement) of human behavior as well as the cognitive profile can influence the perception of the robot and the interaction with it. The strength of this work relies on the fact that we performed a cognitive assessment (i.e., the interaction) with a not anthropomorphic robot, ASTRO, coding basic behaviors to make it appear as social as possible (i.e., by integrating speech synthesis and lighting feedbacks). Indeed, the implemented social behaviors had an impact on the anthropomorphism of the platform. Even if the appearance is not humanlike, the robot’s ability in coaching and motivating the participants affected the perceived consciousness according to the older adults’ personality, i.e., GS2_Consciousness score increased significantly). Additionally, our study involved a group of heterogeneous participants, both from a clinical and a social point of view, allowing us to infer some preliminary conclusions that could be deeply validated in future studies with a larger group of participants.

Regarding RQ1, the analysis reported in Section 4.2 showed a general increase of the final score for all the questionnaires administrated in the T3 phase compared to the scoring obtained in the T1 phase. It highlighted that the interaction with the robot has some positive effects on the person and robot’s perception, especially in the feeling of anxiety,
which significantly decreases after the interaction. The latter finding is also confirmed by our previous work [40].

The analysis of the role of personality suggests that depending on the domain/task, different traits could have an influence. Among the 5 traits assessed by the BFI, mainly Openness to experience, Agreeableness, and Neuroticism were correlated to the perception of the robot, while Openness to experience was the only trait related to the usability of the technology (i.e., the predisposition of using it). Similarly, Extraversion and Neuroticism personality traits mostly influence the modality of the interaction (i.e., total interaction time) with the robot. These results answer to RQ3 (i.e., regarding usability and perception) and partially to RQ5. Specifically, the imaginative and spontaneous traits of the elderly influence the perception of the robot, while the sociality and confidence characterizing the personality have an impact on the interaction itself. In the HRI field, usually, the role of individual traits of the Big Five model is investigated. As an example, Openness to Experience and Agreeableness are taken into account to evaluate the trust aspect of the robotic platform [41–43], while Extraversion and Neuroticism are mostly related to the quality of the interaction [8,44]. While the findings described in [15] relate just Openness to experience with the attitude toward the technology, this study confirms that every facet of personality traits should be taken into account by the robot, to infer different aspects of the interaction.

By analyzing the effect of the cognitive state (i.e., MMSE score) on the interaction performance, we did not find any correlation. This result suggests that the interaction of the participants was not related to their cognitive profile (RQ5). Indeed, the behaviors of the users during the interaction were independent of the cognitive profile, while being highly related to the personality. On the other hand, the cognitive profile of the participants had some effect on the perception of the robotic system (RQ2). As reported in Section 4.5, the highest level of consciousness was attributed to the robotic system by the most impaired participants. This result is aligned with the current trends of adopting toy robots in the later stage of dementia [45,46], as a non-pharmaceutical treatment to stimulate sociality and peace [47].

Regarding the role of emotions felt during the experience, the data recorded at the beginning of the interaction highlight a positive correlation between the capability of feeling strong emotions and the attribution of consciousness to the robotic platform. Additionally, people characterized by positive emotions tended to trust more on the robot’s abilities, after the interaction. From our results, an interesting aspect emerged, namely: People who enjoyed more the interaction trusted more the robot and vice-versa (RQ4). These results are in line with the assessment of the robot’s sensitivity after the interaction, too. It suggests people with a high PANAS score tend to empathize more with the robot, attributing human-like capabilities. Similarly, the presence of non-verbal behaviors (i.e., Shaking head and Nodding) highlighted the tendency of the participants to adopt the same communication they would adopt with another human being. They sympathized with the robot and perceived it as humanlike as if it would understand non-verbal behaviors (RQ5). Similar results were also described in [26], where older people adopted humanlike social conventions with Pepper robot (i.e., touching the hands and the arms of the robot). Our result suggests that even a no-humanlike robot could incentivize social behaviors (see Table 3).

Unfortunately, the consciousness of the robotic platform resulted also positively related to negative affect (NA), after the interaction. It suggests that people who perceived more the robot as conscious stated to be more nervous at the end of the interaction. It may be related to the type of interaction they had with the robot, which may create disenchantment from the previous impression or annoyance. Similarly, results reported that imaginative and curious participants were less surprised by the interaction with the robot. A similar trend is also highlighted by the behavioral parameters manually extrapolated by the videos. On average, the interactions were characterized by neutral emotions, few facial expressions, and a low quantity of motion. It may be related to the nature of the scenario, since carrying
out a cognitive test requires some mental effort, and it may be demanding. As previous works suggest [48], motivation is a key factor in cognitive assessment. The elderly must be motivated while performing it and those who administer the test must be ready to cheer and motivate the elderly at the correct time. In our study, we included some sentences of encouragement, but they may be not enough for keeping the user emotionally enthusiastic.

The main limitation of the presented work relies on the low autonomy of robot behaviors during the cognitive administration. On one hand, the personal factors were manually annotated by the psychologist or directly extrapolated by the answers to the questionnaires. We followed the procedure that is commonly adopted in the cognitive assessment scenario (see Table 2), by including all the necessary tools for assessing the factors of interest. Given the advances in affective computing [49] and automatic personality detection [50] fields, we would like to adopt these solutions to improve the perception capabilities of the robotic platform. It will allow the robot to automatically infer the individual’s traits. On the other hand, cognitive administration has been implemented by employing a Wizard-Of-Oz control. Previous works on cognitive assessment and stimulation adopted available commercial cloud solutions for administrating and scoring cognitive tests automatically [15,16]. The cloud approach may be a possible solution for improving robotic capabilities, but it keeps open the main problem related to the clinical validity of the results. Professional caregivers gain clinical instinct over the experiences, which guides them through the correct assessment of current pathology. This is why every study on socially assistive robotics is usually comprised of two stages: One carried out by the robot and one carried out by a human expert with the scope of validating the previous step (see Table 1). A similar approach has been used also in this work, underling the possibility to integrate the robotic system as a support tool to the professional caregiver. To obtain a reliable framework, robots should gain some expertise that characterizes the professional caregivers, who will be always present in the loop, but that could teach the robot some repetitive tasks to enlarge the screening phase.

6. Conclusions

In conclusion, the results of our study incentivize the usage of a social robot in the assistive scenario. Our analysis suggests that the interaction with the robot is not affected by the cognitive state of the person. It implies that the cognitive test could be administrated indiscriminately at people with different levels of cognitive impairment in the current implemented modality. In the cognitive assessment context, the type of interaction established between the robot and the elderly is influenced by temporal and semi-permanent personal affective disposition. These results give some insights on which elements the robot should take care of to modulate the interaction. To engage the user in a more efficient rehabilitation, the robot should integrate an artificial intelligent-based module to adapt the modality of interaction according to the user’s emotional state. Similarly, the personalization of the cognitive stimulation process should be based on a gradual level of difficulty [51], which could be set based on the interior profile of the participants (i.e., engagement and personality traits).

The main limitations of the proposed work rely on the small sample of participants and the low autonomy of the implemented system. These limitations will guide the future directions of the work.

Author Contributions: Conceptualization, A.S., F.C., and L.F.; methodology, A.S. and G.M.; software, A.S.; validation, A.S. and L.C.; formal analysis, A.S. and G.M.; investigation, L.C. and A.S.; resources, L.F.; data curation, A.S. and G.M.; writing—original draft preparation, A.S. and L.F.; writing—review and editing, A.S., F.C. and L.F.; visualization, A.S.; supervision, L.F. and F.C.; project administration, L.F. and F.C.; funding acquisition, F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was pursued within the CloudIA project (Regione Toscana, POR CREO FESR 2014-2020, CUP 165.24052017.112000015).
Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the joint Ethics Committee of Scuola Superiore Sant’Anna and Scuola Normale Superiore (Delibera 14/2020 of 18 September 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank the director, the medical staff and the residents of the experimental facility where the experimentation took place for their availability.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cross, E.S.; Hortensius, R.; Wykowska, A. From social brains to social robots: Applying neurocognitive insights to human–robot interaction. Philos. Trans. R. Soc. B Biol. Sci. 2019, 374, 20180024. [CrossRef] [PubMed]

2. Rossi, S.; Ferland, F.; Tapus, A. User profiling and behavioral adaptation for HRI: A survey. Pattern Recognit. Lett. 2017, 99, 3–12. [CrossRef]

3. Clabaugh, C.; Mataric, M. Robots for the people, by the people: Personalizing human-machine interaction. Sci. Robot. 2018, 3, eaat7451. [CrossRef]

4. Rudovic, O.; Lee, J.; Dai, M.; Schuller, B.; Picard, R.W. Personalized machine learning for robot perception of affect and engagement in autism therapy. Sci. Robot. 2018, 3, eaa06760. [CrossRef] [PubMed]

5. Nocentini, O.; Fiorini, L.; Acri, G.; Sorrentino, A.; Mancioppi, G.; Cavallo, F. A Survey of Behavioral Models for Social Robots. Robotics 2019, 8, 54. [CrossRef]

6. Robert, L.P.; Alahmad, R.; Estervood, C.; Kim, S.; You, S.; Zhang, Q. A review of personality in human-robot interactions. Ann. Arbor. 2020, 4, 107–212. [CrossRef]

7. Braun, M.; Alt, F. Affective assistants: A matter of states and traits. In Proceedings of the Conference on Human Factors in Computing Systems-Proceedings, Glasgow, UK, 4–9 May 2019.

8. Ivaldi, S.; Lefort, S.; Peters, J.; Chetouani, M.; Provasi, J.; Zibetti, E. Towards Engagement Models that Consider Individual Factors in HRI: On the Relation of Extroversion and Negative Attitude Towards Robots to Gaze and Speech During a Human–Robot Assembly Task: Experiments with the iCub humanoid. Int. J. Soc. Robot. 2017, 9, 63–86. [CrossRef]

9. Anzalone, S.M.; Boucenna, S.; Ivaldi, S.; Chetouani, M. Evaluating the Engagement with Social Robots. Int. J. Soc. Robot. 2015, 7, 465–478. [CrossRef]

10. Khamassi, M.; Velzantzas, G.; Tsitsimis, T.; Tzafestas, C. Robot Fast Adaptation to Changes in Human Engagement During Simulated Dynamic Social Interaction With Active Exploration in Parameterized Reinforcement Learning. IEEE Trans. Cogn. Dev. Syst. 2018, 10, 881–893. [CrossRef]

11. Jain, S.; Thiagarajan, B.; Shi, Z.; Clabaugh, C.; Mataric, M.J. Modeling engagement in long-term, in-home socially assistive robot interventions for children with autism spectrum disorders. Sci. Robot. 2020, 5, eaaz3791. [CrossRef]

12. Hamada, T.; Okubo, H.; Inoue, K.; Maruyama, J.; Onari, H.; Kagawa, Y.; Hashimoto, T. Robot therapy as for recreation for elderly people with dementia - Game recreation using a pet-type robot. In Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN, Munich, Germany, 1–3 August 2008; pp. 174–179. [CrossRef]

13. Tapus, A.; Tapus, C.; Mataric, M.J. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In Proceedings of the 2009 IEEE International Conference on Rehabilitation Robotics, Kyoto, Japan, 23–26 June 2009; pp. 924–929.

14. Fan, J.; Bian, D.; Zheng, Z.; Beucher, L.; Newhouse, P.A.; Mion, L.C.; Sarkar, N. A Robotic Coach Architecture for Elder Care (ROCARe) Based on Multi-User Engagement Models. IEEE Trans. Neural Syst. Rehabilitation Eng. 2017, 25, 1153–1163. [CrossRef] [PubMed]

15. Rossi, S.; Santangelo, G.; Staffa, M.; Varrasi, S.; Conti, D.; Di Nuovo, A. Psychometric Evaluation Supported by a Social Robot: Personality Factors and Technology Acceptance. In Proceedings of the 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Tai’an, China, 27–31 August 2018; pp. 802–807.

16. Di Nuovo, A.; Varrasi, S.; Lucas, A.; Conti, D.; McNamara, J.; Soranzo, A. Assessment of Cognitive skills via Human-robot Interaction and Cloud Computing. J. Bionic Eng. 2019, 16, 526–539. [CrossRef]

17. Desideri, L.; Ottaviani, C.; Malavasi, M.; Di Marzo, R.; Bonifacci, P. Emotional processes in human-robot interaction during brief cognitive testing. Comput. Hum. Behav. 2019, 90, 331–342. [CrossRef]

18. Palestra, G.; Pino, O. Detecting emotions during a memory training assisted by a social robot for individuals with Mild Cognitive Impairment (MCI). Multimedia Tools Appl. 2020, 1–16. [CrossRef]

19. König, A.; Satt, A.; Sorin, A.; Hoory, R.; Toledo-Ronen, O.; Derreumaux, A.; Manera, V.; Verhey, F.R.J.; Aalten, P.; Robert, P.H.; et al. Automatic speech analysis for the assessment of patients with predementia and Alzheimer’s disease. Alzheimer’s Dementia: Diagn. Assess. Dis. Month. 2015, 1, 112–124. [CrossRef]

20. Rossi, S.; Conti, D.; Garramone, F.; Santangelo, G.; Staffa, M.; Varrasi, S.; Di Nuovo, A. The Role of Personality Factors and Empathy in the Acceptance and Performance of a Social Robot for Psychometric Evaluations. Robotics 2020, 9, 39. [CrossRef]
50. Mehta, Y.; Majumder, N.; Gelbukh, A.; Cambria, E. Recent trends in deep learning based personality detection. *Artif. Intell. Rev.* **2020**, *53*, 2313–2339. [CrossRef]

51. Aranha, R.V.; Correa, C.G.; Nunes, F.L.S. Adapting software with Affective Computing: A systematic review. *IEEE Trans. Affect. Comput.* **2019**, *1*. [CrossRef]