2020 29th IEEE International Conference on
Robot and Human Interactive Communication (RO-MAN)
August 31 - September 4, 2020 - Virtual Conference

Multidimensional evaluation of telepresence robot: results from a field trial*

L. Fiorini, *Member, IEEE*, *EMBS* G. Mancioppi, C. Becchimanzi, A. Sorrentino, M. Pistolesi, F. Tosi and F.Cavallo, *Member, IEEE*

Abstract— **The European population is getting older; many elderlies would like to live independently in their homes as long as possible. In this context, a robotic telepresence service could support the frail persons in their homes, empowering their social relationships. In this work, 10 frail elderly were asked to live with a telepresence robot (i.e. Double Robot), through which the formal caregiver could remotely visit and chat with them. A total of 169 days of field-test trial was evaluated before (TO) and after (TF) the tests with a multidimensional framework, including acceptance, usability, and expectation domains. The system was used for 2871 mins, and the results underline good usability- and acceptance- related domains (average score equals to 70.83 at TF) and expectation (average score equal to 67.01 at TF). Results remark that expectation could influence the potential and the real use of the robot. Additionally, a positive trend in the answers was identified between T0 and TF. Indeed, the evaluation of a system should envisage a complex, multidisciplinary and holistic approach, that may influence the success or failure of the robot's purpose, if not properly analysed during the evaluation and design phase.**

I. INTRODUCTION

According to the worldwide projections by 2050 people over 65 will be more than double compared to the children under five. By 2050, the number of people aged 65 and over will also exceed the number of teenagers and young people between the age of 15 and 24 [1]. Aging causes a decline in cognitive and physical abilities. Among such a framework stands out the need for temporary or continuous assistance from a caregiver, namely a family member or a health/social worker. That usually forces the elderly to move from their home to nursing homes, radically changing their everyday life, both from a practical-organizational and a psychological-relational point of view. As pointed out by some scientific evidence [2], the physical environment is crucial for the wellbeing of individuals, mostly for those who are getting older. The majority of the elderly want to live in their own homes, as long as possible. That is because almost all their fundamental actions are carried out there. Moreover, such a scenario is also promoted by policy-makers and health providers to avoid the costly options of institutionalized care [3]. However, one of the consequences of aging is the risk of accidents, such as falls or other domestic injuries, but also problems of isolation or depression and loneliness.

In this context, technology could provide support for the “aging in place” defined as: “the ability to live in one’s own home and community safely, independently, and comfortably, regardless of age, income, or ability level” [4]. A recent systematic literature review highlights the cruciality of developing new technological solutions for the assessment, the treatment, and the monitoring [5]. Recently, telehealth systems, such as remote monitoring services, were developed and tested in real environments. Only during the lasts years, robotics technology has become robust and reliable enough to provide for the employment of long-term evaluations in in-home settings [6]. This approach allows investigating which human factors mostly influence the experience that people have of a particular technology.

Literature review [7] underlines that several domains could influence the impact of a specific technology on someone’s life. Authors underline that the impression of a particular technology could change pre- and after- the trial since users can understand its potential positive impact on their daily life. Indeed, the user establishes a complex relationship with the product-service system, which is based on physical contact, sight, touch, hearing, smell, but also on understanding the functioning and feedback, and on the degree of appreciation of the product [8]. Therefore, the concepts of usage, the intention of use, and expectations are inseparable, especially concerning the interaction with products with a strong social and interactive component, such as assistive robotics. For instance, Rossi et al. [9] showed that the perceived trust of the human in a robot drops drastically when a robot presents behaviours that can lead to severe consequences. Chanseau et al. [10] found out participants’ perception of robots was higher when the robot was performing a task in a semi-autonomous mode.

In this context, this paper presents a robotic telepresence service tested in real houses. The proposed service allowed the social workers to remote visit the frail seniors at their home, increasing the total number of visits performed in a week. After a field test trial, where the users were asked to live with the telepresence robot for at least ten days, the users assessed the service. The proposed multidomain evaluation includes several domains, such as expectations, safety, anxiety, trust, enjoyment, ease of use, the intention of use, and the perceived psychological communication. Two questionnaires were administrated at the beginning and the

* Research was pursued within the SUMA project - Struttura Urbana Multifunzionale Attiva (Progetti Regione Toscana PAR FAS Bando FAS SALUTE – CUP D51B18000740009).

L. Fiorini, G. Mancioppi, A. Sorrentino and F. Cavallo are with the BioRobotics Instutute of Scuola Superiore Sant’Anna, Pontedera (Pisa), Italy and with Department of Excellence in Robotics & AI, Piazza Martiri della Libertà, Pisa, Italy (correspondent author laura.fiorini@santannapisa.it).

C. Becchimanzi, M. Pistolesi and F. Tosi are with the Department of Architecture, Ergonomics & Design Lab (LED), University of Florence, Florence, Italy.

F. Cavallo is with the Department of Industrial Engineering, University of Florence, Florence, 50139, Italy.

978-1-7281-6075-7/20/$31.00 ©2020 IEEE

end of the trial, thus to investigate the following research hypothesis (HYP):

Hyp I) Which factors related to the expectation mainly influence the experience of the robot?

Hyp II) Could the expectation of using the robot influence its usability and acceptance after the use?

Hyp III) How could real experience of the robot modify a person's thoughts?

II. Related Works

Over the last years, people were asked to evaluate the robot after a short and controlled interactions. For instance, within RAMCIP project, elderlies and people suffering from Mild Cognitive Impairment were asked to assess the proposed scenarios from a user experience point of view [11]. Similarly, in the Robot-Era project, the users were asked to evaluate the robots mainly from acceptance and usability point of view after one single interaction [12]. As recently underlined in [13] few papers, attempts to evaluate the robotic system in private homes without the presence of technical staff.

On the contrary, some researches focused on the evaluation of robots in real settings. For instance, De Graaff et al. [6] presented an evaluation of Karotz robot over several domains (i.e. Attitudinal beliefs, Social normative Beliefs, Control Beliefs, Outcome variables) six times during the experimental set-up. However, the authors used a robot that was not able to move in the environment. Conversely, Cesta et al. [14] evaluate the Giraff robot for the telepresence service after one year of usage in a private home as a case study. The authors involved only three subjects (i.e. a couple of older seniors and his son as caregivers) and mainly focused on the evaluation of the expectation-related parameters. Gross et al. [13] evaluate the SYMPARTNER robot which lived for 5 days in 20 elderlies' houses. The evaluation focused mainly on the system's robustness and reliability, such as on the evaluation of the proposed scenarios and the perceived enjoyment. Eventually, Manzi et al. [15] introduced the Kubo robot, inspired by a coffee table that was evaluated in a real apartment with a couple of elderly for 5 days. At the end of the experimental test, the robot was evaluated, using a semi-structured interview, from a technical and user point of view. SERROGA project [13] tested the robot companion with 9 elderly for three days in their home, and their evaluation was mainly focused on instrumental and social-related domains.

In this context CloudIA project (funded by Tuscany Region) aims to provide remote support through a telepresence robot to frail elderly. The social rationale behind this choice came from an exploratory phase conducted with four Tuscan social cooperatives in which the social workers and directors were interviewed. The results a strong need to support the social workers and improve the socialization and the monitoring of people included in the domiciliary assistive program. Therefore, the proposed telepresence service is a potential solution to improve both the quality of elders' life and the services provided to them as underlined by Moyle et al. [16]. Indeed, it can support their communication possibilities, promote social inclusion and increase the real, and perceived security.

III. Material and Methods

The following section introduces the double robot, summarizes the experimental protocol, describes the participants' cohort, such as the evaluation tools, and the data analysis.

A. Double Robot

The Double robot is a trade product designed by Double Robotics (Burlingame, California). Studying an existing commercial robot has been successfully applied to investigate people's desires and expectations and was argued to be favoured over robotic research prototypes, which are often not suitable for long-term studies in real environments. The Double robot is mainly composed of three parts: the robot, the iPad (Apple, California), and the charging station. It is a self-driving, two-wheeled telepresence robot which mounts an iPad. The Robot and iPad are connected through Bluetooth. The Double robot has a proprietary web interface that is used for the telepresence service. After authentication (with username and password), a caregiver could use it to move into the environment and interacting with the user in his home. The installation of the double robot in the home is very plug and play, it just needs the internet connection. Since most of the recruited people did not have an internet connection, a router 4G was installed in their home for the duration of the field test. When the telepresence service is not activated, the Double robot remains at the charging station. Before finish the call, the caregiver send it back at the charging station.

B. Participants

The recruitment process was conducted by Pane & Rose social cooperative, which was involved in the CloudIA project, founded by Tuscany Region. That cooperative, operating in the municipality of Prato, contacted the frail older adults (aged over 60 years old) that lived in their home and that already were beneficiary of home assistant care. At the end of the recruitment process, a total of ten older adults accepted to be involved in the field trial (aged: 75.10 ± 16.48 years old). Four of them were male, and six participants were female. One person was married, five persons were widowed, and the remaining were not married. Two persons lived alone, and the others lived with relatives. As concerns the educational level, five persons attended the elementary school, four persons the middle school, and the remaining had a higher educational level. Before starting the trial, all the participants signed an informed consent form.

C. Experimental protocol

The proposed protocol is divided into three main phases as depicted in Figure 1. In the first part of the experimental protocol, the engineers trained the social workers to the use of the Double robot and to install it in a private home. This phase took place in a laboratory setting, while the other two (i.e. training in situ and field test) took place at the user's private home. Firstly, the caregiver has to install the robot, set-up the wi-fi connection (if necessary), explained the service, and train the user. After that, the field test phases could start. Before starting, the caregiver administrated two

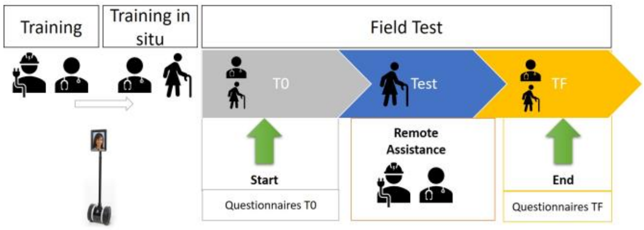

Figure 1. Experimental protocol that included the training of the social operators in laboratory setting, the in-situ training of the experimental participants and the field test where the user was set free to use the Double robot.

questionnaires, and then he left the home. The user was requested to interact with the Double robot without the supervision of the caregiver, thus the users were free to use the robot as they wanted. To avoid overstressing the participants, the caregiver fixed a time slot during which he/she could call the participant (e.g. before lunch, early morning, after lunch) without interrupting their routine. In our set up, the caregiver called the user; if the user accepted the call, the caregiver and the users started to talk, if necessary, the caregiver could move into the environment. Before ending the call, the caregiver put the robot on-charge. At the end of this trial, the caregiver came back to the house to address the same questionnaire he administered at the beginning.

C. *Evaluation tools*

To achieve the scientific goals of this paper, two multidimensional questionnaires were used. Both questionnaires, described below, were administrated at the beginning (T0) and ad the end (TF) of the field test.

Ad-hoc questionnaires: this questionnaire, adapted from the Robot-Era project [12], aims to evaluate several domains related to the usability and acceptance of a certain technology. It was composed of 14 items and it can be divided into the following subdomains:

- *Intention to use* (ITU): it assesses if the user will use the Double in case of need (Q1), if it will help the family's work (Q2), and improves his/her independence (Q3).
- *Anxiety (ANX) and Enjoyment (ENJ)*: it assesses if the user was embarrassed (Q4), and feel nervous (Q5), or had fun (Q6) while using the robot.
- *Trust (TRU):* it assesses if the user will trust the Double's ability to perform telepresence (Q7) and if he/she considers it too invasive for their privacy (Q8).
- *Perceived Ease of Use (PEU):* it investigates if users find the system easy to use (Q9), if they understand how to correctly handle the call (Q10), if they find the Double interface easy to use (Q11), and if it was easy to talk to the caregiver (Q12).
- *(only for the caregiver) Perceived Ease of Use (PEU)* to investigate if drive the double into the house (Q13), and put it back into the charger was easy (Q14).

All these items were evaluated on a 5-point Likert scale.

Expectation: similarly to [14] we used a questionnaire composed of 13 items to evaluate the expectation and it can be divided in likeability, perceived intelligence and perceived safety sub-scales. Participants were requested to evaluate on a 5 point likert scale if they strongly agree/strongy disagree to each item's statement (*"I think that"* + *item*). In the following, for each subscale we detail the items included:

- Likeability (LIK), includes the item 3 (Afraid to not good at use), item 8 (Perceived an increased distance of the caregiver during the service), and item 12 (consider the robot as a negative component of the communication).
- Perceived Intelligence (PEI), includes the item 4 (Increased independence), Item 5 (reduced need of presence), item 6 (help in taking care), item 7 (Difficult in the communication), item 9 (improved contacts) and item 13 (difficulty to interact).
- Perceived Safety (PES), includes item 1 (privacy concern), item 2 (be afraid of the robot), item 10 (increased security), item 11 (support in an emergency).

D. *Data analysis*

Firstly, the interclass correlation coefficient (ICC) and the Cronbach's alpha were calculated to assess the questionnaires' reliability, particularly the internal consistency. Afterwards, the mean average for each item was computed, then to obtain an overall score of the questionnaires, the item score was rescaled between 0-100. Negative question scores were inverted.

Non-parametric tests were applied to compare different conditions and users. Non-parametric statistics were necessary because of the sample size and data distribution. As for the Hyp I, the Spearman correlation coefficient (RHO) was computed to observe the relationship between expectation and ad-hoc questionnaires' scores, both at T0. As concerns the Hyp II, RHO was computed between expectation and ad-hoc questionnaires' scores at T0 and TF. Both analyses were computed at items and sub-scales levels.

Conclusively, for Hyp III, the Wilcoxon tests for paired samples were used to compare the expectation and the ad-hoc questionnaire scores at the baseline (T0) and after the experience with the robot (TF). Furthermore, Mann-Whitney U tests have been applied to compare demographic variables (age, gender, educational level, and the number of relatives).

IV. Results

At the end of the field trials, which lasted 169 days in total, a total of 105 telepresence services were delivered. The average duration of a single service was equal to 29 mins for a total of 2871 mins of usage. Two subjects did not answer to expectation questionnaire at T0 and TF. Therefore, only the comparative analysis was conducted on 8 subjects for the expectation questionnaire.

The reliability of the questionnaires, calculated on the T0 results, has been confirmed by ICC and the Cronbach's

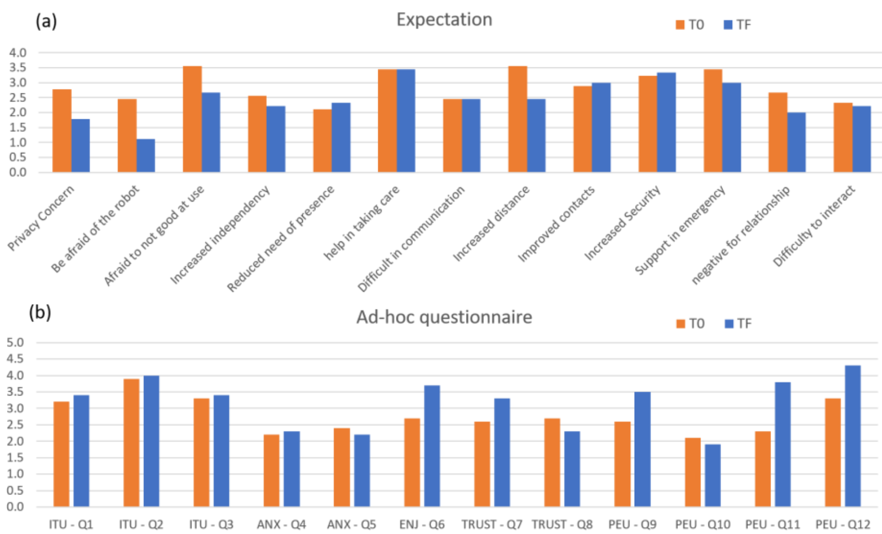

Figure 2. a) Average answer to each items of the expectation questionnaire at T0 (orange bars) and at TF (blu bars); b) Average answer to each items of the ad-hoc questionnaire at T0 (orange bars) and at TF (blu bars).

alpha. ICC values between 0.40 and 0.75 can be considered as good [17]. For both the questionnaires the ICCs were higher than 0.4 (0.49 and 0.73 for the ad-hoc and the expectation questionnaires respectively). Cronbach's alphas were higher than 0.60 for both the questionnaires (0.62 for the ad-hoc questionnaire and 0.76 for the expectation questionnaire). Such values are considered acceptable for short instruments with a small number of items [18], [19]. In conclusion, both of them could be considered reliable.

Regarding the results of both questionnaires, Figure 2 reports the mean value computed for each item. As for the total score, the participant's scores were added together and converted to a new score ranging from 0-100 (negative items were inverted). The mean overall score for the ad-hoc questionnaires (12 items) was 61.17 (SD=12.30) at T0 and 70.83 (SD=9.79) at TF. At T0 one subject was very critical about the service (total score = 38.33) but at the end of the trial, he changed his mind rating the experience as positive (total score =70). At TF, 5 subjects rated the service >70, and 4 subjects between 60 and 70. Concerning the expectation questionnaire, the overall score at T0 was equal to 63.08 (SD=6.53) and at TF was equal to 67.01 (SD=10.35). At TF, 4 subjects rated the service >70 and 3 subjects between 60 and 70.

Correlations between the results of the questionnaires have been calculated. We always considered the expectation at T0, while we used the ad-hoc questionnaire's outcomes both at T0 and TF to investigate Hyp I and Hyp II respectively (see Figure 3). This analysis was performed at items and sub-scale levels. Regarding the correlation between expectation at T0 and the ad-hoc questionnaire at T0 is possible to observe a significant correlation between expectation item 2 (are you afraid for your privacy) and item ANX-Q5 [RHO = 0.765]; expectation item 3 (be afraid of the robot) and item ANX-Q4 [RHO = 0.819]; expectation item 4 (increased independency) and item ANX-Q5 [RHO = -0.673]. The expectation item 7 (difficult in the communication) and expectation item 8 (increased distance) are strongly both correlated with item TRUST-Q7 (I will trust the Double's ability to perform telepresence), [respectively RHO = 0.841 and RHO =0.766]. Other correlations are between expectation item 9 (improved contacts) and item PEU-Q11 [RHO = 0.797]; expectation item 13 (difficulty to interact) and item ANX-Q5, RHO = -0.714. For what concern the correlation with the sub-scales, the PEI and ITU-Q3 are strongly correlated [RHO = 0.836] such as the PES and item ANX-Q5 [RHO = -0.794].

As for Hyp II, concerning the correlation between expectation at T0 and ad hoc questionnaire at TF is possible to observe correlation between: expectation item 4 (increased independency) and item PEU-Q9 [RHO = 0.894]; expectation item 6 (help in taking care) and items ITU-Q2 and ITU-Q3 [RHO = 1.000 and RHO = 0.800 respectively]. The expectation item 7 (difficult in communication) is correlated with items TRUST Q7 and PEU Q-12 [RHO = -0.706 and RHO = 0.716 respectively]; the expectation item 10 (increased security) is correlated with PEU Q-11 and PEU Q-12 [RHO = 0.854 and RHO = 0.834 respectively]; the expectation item 11 (support in emergency) is correlated with either items ITU Q2, ITU Q3, and ENJ Q6 [respectively RHO = 0.921, RHO = 0.746, and RHO = 0.713]. As concern the correlation at sub-scale level the PEI is correlated with items ITU Q2, ITU Q3, ENJ Q6, and PEU Q9 [RHO = 0.889, RHO = 0.759, RHO = 0.678, and RHO = 0.805 respectively]; and PES is correlated with items ENJ Q6 and PEU Q9 [RHO = 0.706 and RHO = 0.781 respectively].

As concern the Hyp III, which is related to the analysis of the effect of the robot's usage, the results show no significant statistical differences between T0 (before the use of the robot) and TF (after the use of the robot) except for the item 2 of the expectation questionnaire (does the robot scare you?) and the PEU-Q11 (do you find the interface easy to use?). Therefore, the experience with the robot positively affects these two parameters which significantly differ after the usage of the Double. Some other positive trends can be observed in Figure 2. Regarding the expectation

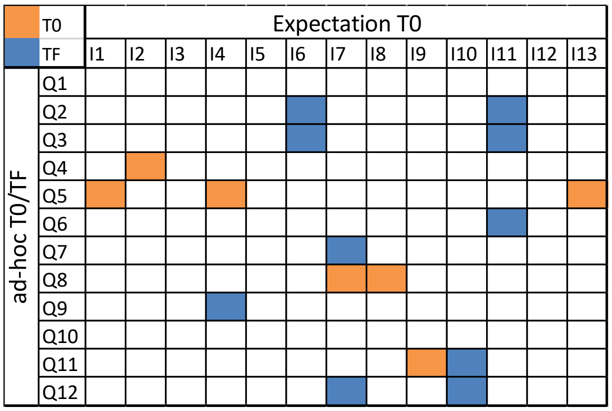

Figure 3. Summary of the significant correlation analysis ($p<0.05$) at item level between the Expectation questionnaire at T0 and ad-hoc questionnaire at T0 (orange) and TF (blue).

questionnaire (Figure 2.a) Item 1 (Privacy Concern), Item 3 (Afraid to not good at use), Item 8 (increased distance) such as Item 12 (negative for the relationship respectively) show a positive trend. Indeed, at the end of the field test (TF) the average values decrease, which means that they were evaluated in a slightly positive way (these are all negative questions). Figure 2.b reports the trend of the other questionnaire, particularly, it is worth noticing that ENJ-Q6, TRUST-Q7, PEU-Q9, and PEU-Q12 reports a positive trend after the use. No significant statistical results have been identified comparing expectation and the ad-hoc questionnaire's score, both at T0 and TF, to demographic characteristics of our sample: age, gender, educational level, number of relatives.

V. Discussion

This paper aims to present and discuss how human factors could influence the potential and the real experience of the telepresence services (Hyp I, Hyp II respectively), but also how a real field test in the private home could change the mind of recruited subjects (Hyp III). Two multidimensional questionnaires were administered before and after the test. Both overall scores increased after the field tests, which could suggest that a real experimental test could be useful to "better understand" the technology. The correlations analysis shows how expectations, before to use the robot, are related to what people suppose about robot's usability, acceptance, and usefulness (expectation T0 / ad-hoc T0). Moreover, the analysis of correlations shows how such preconceptions are related to the final the ad-hoc questionnaire's outcomes (expectation T0 / ad-hoc TF) (Figure 3).

Regarding the Hyp I, the correlations coefficients between expectation at T0 and the ad hoc questionnaires at T0 denote that subjects who expected the robot negatively affects their privacy (expectation item 1), their independence (expectation item 6), and their safety (PES) felt also more embarrassed (ANX-Q4) and nervous (ANX-Q5) using the robot. An odd data came out from the relation between the ANX-Q5 and Item 13 of expectation questionnaires. It seems that during the interaction with the robot the less he/she thinks the robot could negatively affect the rate in which the caregiver visits them or the robot could make the communication difficult. This finding should be carefully studied, even though an explanation might be related to the way the questions have been proposed. Probably this question has been poorly understood. On the other hand, subjects who had a higher overall expectation (expectation total score) regarding the use of the robot showed a higher level of trust (TRUST-Q7) in robot ability and presumed that the robot could ameliorate their independence (ITU-Q3). ITU-Q3 is also related to a higher perception of the robot's intelligence (PEI). On the contrary, subjects who believed that the double robot was too invasive (ITU-Q8) showed also the high expectation of high difficulty in the communication (expectation item 7) using the Double. Besides, the perception of an easy and intuitive interface (PEU-Q11) correlates positively to the expectation of a stable connection with the caregiver (expectation item 9).

Concerning the Hyp II, some useful information can be gathered by the correlation between expectation scores at T0 and the ad-hoc questionnaire scores after the experience with the robot (TF). In this case, subjects who considered that the robot would permit the caregiver to be more present and close to themselves (expectation item 6), and also that the robot could be useful in cases of emergency (expectation item 11), show the higher score at TF believing the robot could foster their autonomy (ITU-Q3), and that it could be useful for the family system (ITU-Q2). In general, the idea that the robot could foster someone's autonomy (ITU-Q3) is related to higher total expectation at T0 (expectation total score). Such items (ITU-Q2 and ITU-Q3) are also positively related to the overall perceived intelligence of the robot (PEI). Moreover, highest enjoyment scores (ENJ-Q6) at TF are related to the perception of the robot as an intelligent and safe agent (PEI and PES) and to its perceived usefulness in case of emergency (expectation item 11). Such score (ENJ-Q6) is also correlated to the overall expectation score towards at T0. This result is also aligned with the state of the art that underlines the important role that enjoyment plays toward a new technology [7]. Along with the aforementioned results, is present a negative correlation between the scores of subjects who believe the robot could make the communication difficult and the TF scores about the robot's telepresence ability. Higher scores PEI and PES, correlate to how subjects found the robot easy to use (ITU-Q9). A positive correlation is also present between the ITU-Q9 and the overall expectation toward the robot at T0. Furthermore, the higher expectation in believing the robot capable of making subjects safer (expectation item 10) is related to how subjects perceived the interface easy to use (ITU-Q11) and to how the robot could foster the possibility to interact with the caregiver (ITU-Q12). Another similar research underlines that people tend to positively evaluate technology as they become familiar with it [6].

Correlational data are not the only data that gave to us some useful information. The variance analyses, performed using the Wilcoxon test for paired samples, permitted to study how expectation and the usability/acceptance perception changed over the interaction with the robot. In particular, the expectation item 2 (be afraid of the robot) is a statistically significant difference if we compare T0 to TF. Graphics reported in Figure 1 show us a decrement of the degree of fear towards the robot after its usage. Other positive trends, even though not statistically significant, are detectable. Particularly, after the field trials, the subject's concerns regarding their privacy decrease, and also the perception to be not able to use the robot. Furthermore, the thoughts concerning the increasing distance made by the robot, as well as the negative impact on the relationship with the caregiver have dropped after the experience with the robot. Lastly, the idea that the robot increases the security slightly improve.

As for Hyp III, the Wilcoxon test for paired samples on ad-hoc questionnaires show us statistically significant differences between item 11 (the interface usability) at T0 and TF. Such parameters arise after the robot's usage. The same trend can be observed along with all the PEU sub-scale items, apart from item 10. This trend suggests that perceived easiness related to the robot's use and interface may increase through the experience. Furthermore, the TRUST subscale shows an interesting pattern. These data encourage to believe that making real experience with the robot could foster the trust, and increase the perceived comfort during the interaction. Another interestingly, but not statistically significant, data concern the enjoyment (ENJ-Q6). After using the robot, people declare to enjoy more than they expected [7]. Based on our small sample, it seems that there is no relationship between gender, age, educational level, and family composition in any measurements.

VI. Conclusion

Conclusively, we can suppose that prejudice on the robot (expectation T0) could affect the perceived robot's usability/acceptance (ad-hoc questionnaire T0), as well as the real robot's usability and acceptance (ad-hoc questionnaire TF). In particular, a positive *a priori* attitude is related to more trust towards the robot, more enjoyment in using it, and with the idea that the robot is easy to use. On the other hand, negative bias towards the robot (i.e. to be afraid of the robot) could increase the subjects' anxiety and the tension during the interaction. Moreover, the use of the robot itself could improve some parameters. After using the robot, subjects generally feel more comfortable, confident, and less scared about it. Expectation and prejudice seem to play a key role, not only in determining the perceived usability/acceptance of a robot but also concerning the real usability and acceptance after the interaction with the robot. As a robot's designer and developer, we need to meet the subjects' needs and personality traits fostering trust. Moreover, to promote a positive attitude towards the robot, we should also intervene on the end-users themselves, reducing the gap between laboratory and real use. For instance, seminars and training sessions on robots should be organized to enhance user expectations and to reduce that gap. This type of methodology combined with the familiarization with the robot might promote the cooperation between the user and the machine.

Acknowledgement

The authors would like to thank Lara Toccafondi and Carolina Faini from Pane & Rose Social Cooperative for the recruitment process such as the experimental phase. Special thanks also to all people directly involved in the experimentation.

References

[1] WHO, "Health workforce: Data and statistics," 12-Mar-2018. [Online]. Available: http://www.euro.who.int/en/health-topics/Health-systems/health-workforce/data-and-statistics. [Accessed: 12-Mar-2018].

[2] L. Fiorini *et al.*, "Assistive robots to improve the independent living of older persons: results from a needs study," *Disabil. Rehabil. Assist. Technol.*, pp. 1–11, Jul. 2019.

[3] M. A. Williams, R. L. Soiza, A. M. E. Jenkinson, and A. Stewart, "Exercising with Computers in Later Life (EXCELL) - Pilot and feasibility study of the acceptability of the Nintendo WiiFit in community-dwelling fallers," *BMC Res. Notes*, vol. 3, no. 1, pp. 1–8, Sep. 2010.

[4] "CDC - Healthy Places - Healthy Places Terminology." [Online]. Available: https://www.cdc.gov/healthyplaces/terminology.htm. [Accessed: 23-Mar-2020].

[5] G. Mancioppi, L. Fiorini, M. Timpano Sportiello, F. Cavallo, M. T. Sportiello, and F. Cavallo, "Novel Technological Solutions for Assessment, Treatment, and Assistance in Mild Cognitive Impairment: A Systematic Review," *Front. Neuroinform.*, vol. 13, p. 58, Aug. 2019.

[6] M. M. A. De Graaf, S. Ben Allouch, and J. A. G. M. Van Dijk, "Long-term evaluation of a social Robot in Real Homes," in *AISB 2014 - 50th Annual Convention of the AISB*, 2014, vol. 17, no. 3, pp. 462–491.

[7] A. Tsertsidis, E. Kolkowska, and K. Hedström, "Factors influencing seniors' acceptance of technology for ageing in place in the post-implementation stage: A literature review," *Int. J. Med. Inform.*, vol. 129, no. June, pp. 324–333, 2019.

[8] D. A. Norman, *Emotional design*. Cambridge: Basic Books, 2003.

[9] A. Rossi, K. Dautenhahn, K. L. Koay, and M. L. Walters, "How the Timing and Magnitude of Robot Errors Influence Peoples' Trust of Robots in an Emergency Scenario," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10652 LNAI, pp. 42–52.

[10] A. Chanseau, K. Dautenhahn, K. L. Koay, M. L. Walters, G. Lakatos, and M. Salem, "How does peoples' perception of control depend on the criticality of a task performed by a robot," *Paladyn*, vol. 10, no. 1, pp. 380–400, 2020.

[11] J. Gerłowska *et al.*, "Assessment of perceived attractiveness, usability, and societal impact of a multimodal Robotic Assistant for aging patients with memory impairments," *Front. Neurol.*, vol. 9, no. JUN, pp. 1–13, 2018.

[12] F. Cavallo *et al.*, "Acceptance of Robot-Era system: results of robotic services in smart environments with older adults (Preprint)," *J. Med. Internet Res.*, vol. 1, Feb. 2018.

[13] H. Gross, A. Scheidig, M. Steffen, B. Sch, C. Fricke, and S. Meyer, "Living with a Mobile Companion Robot in your Own Apartment - Final Implementation and Results of a 20-Weeks Field Study with 20 Seniors *," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 2253–2259.

[14] A. Cesta, G. Cortellessa, A. Orlandini, and L. Tiberio, "Long-Term Evaluation of a Telepresence Robot for the Elderly: Methodology and Ecological Case Study," *Int. J. Soc. Robot.*, vol. 8, no. 3, pp. 421–441, 2016.

[15] A. Manzi *et al.*, "Design of a cloud robotic system to support senior citizens: the KuBo experience," *Auton. Robots*, vol. 41, no. 3, pp. 699–709, 2017.

[16] W. Moyle, U. Arnautovska, T. Ownsworth, and C. Jones, "Potential of telepresence robots to enhance social connectedness in older adults with dementia: An integrative review of feasibility," *International Psychogeriatrics*, vol. 29, no. 12. Cambridge University Press, pp. 1951–1964, 01-Dec-2017.

[17] M. Shoukri and C. Pause, "Statistical analysis of reliability measurements," in *Statistical Methods for Health Sciences*, United States: CRC Press, 1999, pp. 20–28.

[18] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, and U. Saddle River Boston Columbus San Francisco New York Indianapolis London Toronto Sydney Singapore Tokyo Montreal Dubai Madrid Hong Kong Mexico City Munich Paris Amsterdam Cape Town, "MULTIVARIATE DATA ANALYSIS A Global Perspective."

[19] K. Van Der Horst, S. Kremers, I. Ferreira, A. Singh, A. Oenema, and J. Brug, "Perceived parenting style and practices and the consumption of sugar-sweetened beverages by adolescents."