SERMAS

Socially-acceptable Extended Reality Models and Systems

D4.2 SERMAS Human mechanisms in interaction processes

30 December 2023

Funded by the European Union

DELIVERABLE INFORMATION	
Deliverable leader	UNIMORE
Document type	
Document code	D4.2
Document name	Human mechanisms in interaction processes
Work Package / Task	WP4 / T4.2 Human mechanisms in interaction processes
Delivery Date (DoA)	30/12/2023
Actual Delivery Date	
Reviewers	KCL

DELIVERABLE HISTORY			
Date	**Version**	**Author**	**Summary of main changes**
20/10/2023	V0.1	UNIMORE	Drafting the TOC and general content
15/12/2023	V1.0	UNIMORE	Filling the sections, adding content and details; minor changes in the titles of the sections; collecting contributions from the other partners.
20/12/2023	V1.1	UNIMORE	Preparing final version for submission; collecting contributions from the other partners.

DISSEMINATION LEVEL		
PU	Public	
SEN	Sensitive, limited under the conditions of the Grant Agreement	
R-UE/EU-R	EU RESTRICTED under the Commission Decision No2015/444	
C-UE/EU-C	EU CONFIDENTIAL under the Commission Decision No2015/444	
S-UE/EU-S	EU SECRET under the Commission Decision No2015/444	

SERMAS partners

Disclaimer

This project has received funding from the Horizon Europe programme under the Grant Agreement 101070351.

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the European Commission can be held responsible for them.

SERMAS • Grant Agreement: 101070351 • 2022 – 2025 | Duration: 36 months

Topic: HORIZON-CL4-2021-HUMAN-01-13

Public Executive Summary

The present deliverable:

- describes the user study we designed to assess how users' performance and subjective evaluation of the interaction (i.e., of the task, robot, and situation) are affected by the introduction of different interaction modalities (reactive and proactive) compared to a condition in which users are required to perform a challenging assembly task by their own;
- presents the results of the analyses of the performance and of the subjective evaluations provided by the users;
– discusses the relevance of the results and the future research extensions.

Table of contents

List of Figures

List of Tables

1. Introduction

The SERMAS project is aimed at developing XR systems that are not just innovative and functionally complex, but also provide an experience that satisfies the goals and needs of the user, are transparent, safe, secure, explainable, and trusted by the user and are in compliance with the social context in which they are used. In other words, such systems need to be socially accepted by users. The concept of social acceptance is crucial in the SERMAS project. Since in the literature different definitions can be found that underline the contribution to acceptance of several factors, such as robot physical appearance and behavior [e.g. 1], here we focus on the assessment of how the interactive behavior of robotic systems may affect how users perceive them and the interaction. In general term, we are interested in the interaction between a human user and a virtual agent, which in principle could be a virtual avatar or a physical robot. In the current work, we focus our analysis on the interaction occurring between humans and robots during human-robot collaboration tasks. We focus on robotic systems because they are increasingly pervasive in our daily lives, finding applications in diverse environments such as schools, hospitals, and hotels, where they provide relevant services to users or assist people in collaborative operations. For instance, in line with one of the SERMAS pilots, a receptionist robot may need to help and guide visitors, it may provide them with instructions and answer their questions. In such scenarios, robots are not merely asked to execute a task, but to collaborate, that is to act jointly with humans, a requirement that is crucial for SERMAS.

Efficient collaboration requires robots to be socially accepted by users. The concept of social acceptance is crucial for the SERMAS project and research efforts are being devoted to understanding which are the factors that enhance the social acceptance of collaborative robots. The starting point is represented by the characterization of human mechanisms in human-robot interaction (HRI) processes, to gain information that can then be sued to tailor robot behaviors to human needs. To achieve this goal, one can resort to the mature methods developed in the psychology domain for studying human joint collaborations and adapt them to the HRI context.

In order to better understand human behavior in HRI, we have designed a human-robot collaborative task, where a user is asked to build a tower with toy bricks. The user must stack the bricks in colored layers following an order briefly presented at the start. This task challenges the user in two ways: cognitively, as the color order must be memorized in advance, and physically, as the building blocks are initially positioned far from the tower's building site. A small-sized mobile robot can assist the user at with two *levels of collaboration*: (i) *reactive*, in which the robot responds to explicit requests of the user, and (ii) *proactive*, in which the robot autonomously and proactively takes actions to help the user. We designed a collaborative task that is sufficiently simplified to facilitate effective problem structuring, yet it incorporates a diverse range of interaction methods essential for conducting a meaningful analysis. Furthermore, the considered task is also exemplary of real-life human-robot collaborations, e.g. in the domestic context (where collaborative tasks with a robot can reduce the sense of loneliness), healthcare (where robots can help clinicians with medical operations), home assistance (where robots can bring objects to elderly people with mobility issues, for instance).

For this kind of service, it is crucial that the robot is well-perceived and accepted by humans. Thus, the objective of the user study we conducted was to answer the following fundamental research questions:

1. What is the level of collaboration that humans prefer in an HRI collaborative task?
2. What is the impact of the robot's behavior on the human perception of the task and of the robot?
3. Do humans wish to collaborate with a robot trading longer executions for a better experience?

The current study addresses these questions to offer valuable guidance on the behaviors that robots should exhibit to be seen as beneficial and gain user acceptance. In addition to assessing task performance, we assess the users' perception of the robot, the task, and the feelings they experienced during the experiment execution, focusing on cognitive (belief), emotional (feelings), and behavioral (intentions) components of attitudes.

2. Collaboration in HRI

Early research in HRI focused on the functional aspects of robots engaging with humans. However, technological improvements have allowed to design robots capable of engaging in meaningful interactions with the user. This aspect is rapidly becoming crucial, thanks to the dramatic improvement of conversational tools like ChatGPT [2], which allows fluency in HRI inspired by human-human communication [3].

Because of the huge impact of HRI in modern society, research put emphasis on investigating the psychological consequences of interactions between humans and robots [4]. Therefore, the design of successful HRI paradigms encompasses two key dimensions: functional and social perceptions. The functional aspect gauges whether the robot effectively fulfills its service task, while the social dimension assesses the affability and friendliness exhibited by the service robots when engaging with users. As regards social perceptions, one of the most investigated aspects is acceptance, which is a central prerequisite for robots' adoption. Acceptance is commonly defined as the intention to use (or interact with) a robot [5], indeed a lack of acceptance could lead to people's resistance to robots, hindering their introduction in everyday life [6]. To assess human attitude and intention to use technology, the Technology Acceptance Model (TAM) is a very established tool [7] largely in use in a wide variety of different contexts. It has also been already extended to interactive technologies and service robots [8]. The main idea of TAM posits that the intention to employ technology relies on two core factors: the perceived usefulness and the perceived ease of use. The former involves evaluating the anticipated benefits of the technology, the latter pertains to users' confidence in possessing the requisite skills and resources for successfully utilizing a technology. In addition to acceptance, the review in [5] seeks to estimate people's attitudes toward, trust in, and anxiety associated with social robots. Attitude is a multifaceted construct that encompasses people's thoughts and cognitive evaluations (cognition), their feelings and emotions (affect), and observable or self-reported behaviors (behavior) towards a system [9]. Concerning anxiety, a number of studies have provided evidence that anxiety shapes user's intention to use robot, the quality of interaction, and people's behavior during HRI. Trust determines how people assess the reliability of actions and behaviors and form an emotional connection that enforces the expectation in which the interacting robot invests in a common goal [10]. The trust that individuals place in robots is susceptible to influence based on their assessment of the robot's abilities, which can be contingent on factors related to both humans and robots. For instance, it can be impacted by elements like the robot's appearance, degree of autonomy, and range of functions [5,10]. Additionally, users' trust may also be shaped by their individual personalities, self-assurance, and previous experience with robots. With the aim of integrating acceptance and trust in a unified model, the study in [11] investigated the prediction of users' intentions to interact with and to use technology, by integrating different theories from trust literature and beliefs from the theory of planned behavior to TAM. The result is an enhanced trustworthiness beliefs model for robot acceptance that integrates the three theoretical perspectives (trust, beliefs, and acceptance) to enhance the understanding of psychological processes involved in HRI and robot adoption. Their study focused on the cascade from beliefs over attitudes to behavioral intentions to better understand the psychological processes associated with HRI and to facilitate a positive integration of robots in public and private life.

2.1. Interaction modalities: reactive and proactive robot behavior

In humans, collaboration is facilitated by communication. Through explicit communication information is communicated via speech or codified gestures, in a deliberate way so that a person can explicitly ask for help. However, an action or behavior can act as a message, and information is inferred from that action or behavior by another human or agent. We may, for instance, understand that somebody needs help by looking at how this person moves and inspects the environment. To enhance acceptance of robots and improve their social perception, robots should be provided with the ability to react to explicit requests (reactive behavior) but also to anticipate the user's request based on his/her implicit communication, actions, or task status (proactive behavior). Proactive robots are required to understand what is communicated by the human body language prior to approaching a human. Several strategies for generating proactive robot behavior have been addressed in the literature, as reviewed in [12]. However, a large part of this literature in this regard is focused on understanding human intention to interact, for example from posture [13], gaze [14], proxemics [15], or utterances [16]. We believe it is important to investigate how humans perceive proactive behavior by the robot, compared to reactive behavior, where interaction is initiated by the human subject. The user study described in the following section was designed with this aim.

3. Assessment of human-robot collaboration

2.1 User study

The user study was aimed at comparing the effects of different interaction modalities adopted by a robot on user's behavior and perceptions. Specifically, we assessed how users perceive a proactive robot compared to a reactive robot and which is their preferred interaction modality.

2.1.1 Participants

We tested 18 adults (13 males, 5 females, average age 33.28 years, standard deviation 8.6 years) who volunteered to participate in the study. Only 7 participants reported previous experience with autonomous drones, robotic exoskeletons, research robots, and manipulators, but no one with ground mobile robots.
Prior to the beginning of the experiment, all participants signed a consent form stating their voluntary involvement in the experiment, which could be interrupted for any reason at any time. The user study was approved by the SUPSI Ethical Committee.

3.1.2 Task and experimental setup

Participants were tested in a single session. They were required to build a construction using LEGO bricks, on a table, under three experimental conditions. In the passive condition (PA), they performed the task on their own, while in the reactive (R) and proactive (PR) conditions, they performed it in collaboration with the robot. The structure they had to assemble consisted of a small tower composed of 7 layers, each of a different color. The desired pattern was shown at the start of the task for 30 s after which it was covered, and the user had to reproduce it by memory. Each level had to be composed of 4 bricks of the same color with the shape of each level left to the user's will. We used LEGO DUPLO bricks having a dimension of $64 \times 32 \times 24$ mm.
During task performance, the participant sat on a chair at one side of the table and had direct access to only 3 of the 4 bricks needed for completing each layer (see Figure 1). The remaining 7 bricks, one for each color, were placed on the

other side of the table. To complete the task the participant had to use all the bricks at his/her disposal. The presence of a layer of the wrong color was defined as an error and was recorded.

In the passive (PA) condition, participants performed the task on their own without receiving any help from the robot, which stayed still in the middle of the table. This condition served as a baseline to quantify the task difficulty and effort, both physical and mental, needed to accomplish it. In the other two conditions, the robot could assist the user in two ways. In the first one, it could bring (one by one) to the participant a brick required to complete the task (i.e. the LEGO pieces placed on the other side of the table). This aid action could be triggered upon a user's explicit command (reactive condition, R) or autonomously by the robot itself (proactive condition, PR). The second way of assistance was based on audio and visual feedback, and consisted of the following advices provided to the user: (i) color suggestion, according to which the robot communicated the current layer's color to build, by lighting its LEDs with the corresponding color for a few seconds and pronouncing the color name; and (ii) error feedback: when the user assembled a piece of the wrong color, the robot displays a disapproving behavior by flashing the LEDs with red color and pronouncing the word "Error". All the explicit requests of the user happened through vocal commands and the user could ask simple questions to trigger the robot's reaction, as for instance, "Can you bring me the yellow brick?" or "Which color should I use now?". In the proactive condition, the robot behaved autonomously. It did not wait for the user's explicit command to start the pick-and-place motion: once the participant assembled the first brick of the correct color for a new layer, the robot directly picked up the fourth LEGO brick required by the user. Furthermore, the robot could autonomously notify the user of any possible mistake during the building process as soon as it was committed. As in the reactive condition, the user could directly ask the robot for a color suggestion if unsure.

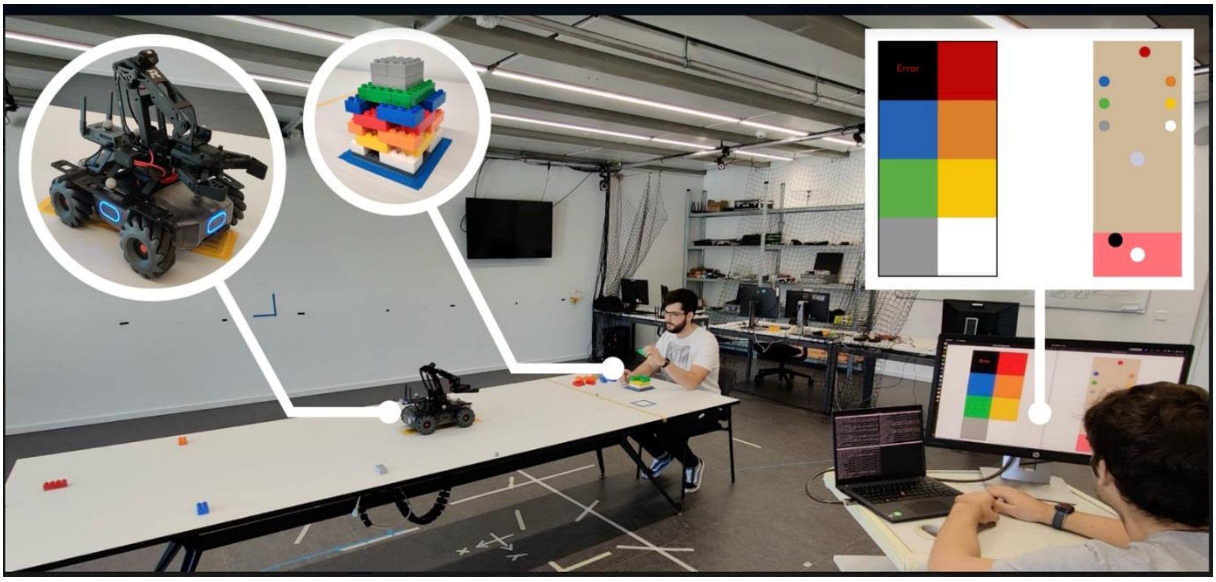

Figure 1: Setup used in the user study

3.1.3 Robotic platform

We used a DJI RoboMaster EP Core, a small-size mobile robot with a length of 0.32 m, a width of 0.24 m, and a height of 0.27 m. This robot is equipped with an omnidirectional wheeled base, a two-degrees-of-freedom arm with a gripper, an onboard RGB camera, a speaker, a microphone, and RGB LEDs. The payload of the arm is about 0.3 kg, providing the robot with the manipulation capabilities required by our collaboration task. The speaker and the LEDs enable the robot to perform non-physical HRI in the form of verbal and visual communication. The relatively small size of the platform allows safe interaction with humans. Furthermore, its nice appearance and smooth motion favor its acceptance with humans and help to increase the trust perceived by users during the collaboration task. The RoboMaster can move its base at a maximum linear and rotational speed of about 3 m/s and 10 rad/s, respectively. However, to display a gentle behavior to the user, these values have been bounded to 2 m/s and 5 rad/s, respectively. The robot's absolute position is controlled using feedback from an external motion capture system. The whole control infrastructure is implemented within the ROS2 framework building on top of a lower-level control node.

3.1.4 Implementation of interaction modalities

To simplify the logic algorithm behind the robot's reactions, we used the Wizard-of-Oz paradigm according to which an operator controls the robot's behavior. The operator sat in the same room as the participant, behind a desk, at about 3 m from the table where the task took place. The Wizard-of-Oz operations were discrete, quiet, and realized through a laptop touchpad. Participants were induced to perceive the robot as an autonomous agent and were told that the operator's presence was required for supervision purposes, e.g. for logging the execution time of the experiment and collecting the errors committed during the execution of the experiment. To simplify the wizard's operations, the robot was controlled through two GUIs. One GUI, shown in the top right box in Figure 1, was dedicated to executing the pick-and-place of the bricks and the other was responsible for the visual and audio feedback. The former was designed to resemble the actual table used in the real setup; the colored circles in the top half showed the bricks' position, the light gray circle in the middle was the robot's idle position, whereas the black circle in the bottom was the brick placing position; the white position indicated the place where the tower was assembled. The robot reaction was triggered by clicking on either one of the brick circles, after which the robot moved towards the specified piece, grasped it, placed it in the placing zone, and then returned to the table center facing the user. The second GUI, dedicated to the execution of audio and light commands, is shown on the left in the top right box in Figure 1. By clicking on the colored square, the robot expressed the corresponding color suggestion or error feedback.

3.1.5 Experimental procedure

Each participant was asked to repeat the task 3 times under each one of the three experimental conditions (PA, R and PR) described above. The construction pattern changed in each session, but it was kept constant across participants in each condition. Participants were divided into 3 groups (i.e. 6 participants per group) who experienced different orders of presentation of the 3 experimental conditions (passive PA, reactive R, and proactive PR interaction conditions), following a Latin square design, i.e. PA-R-PR, R-PR-PA, and PR-PA-R. After each session, participants were asked to fill out a questionnaire (see Section 3.1.7).

3.1.6 Performance measures

During the experiments, the following performance measures were collected:

- time to complete the task, in seconds;
- number of user errors upon task completion;
- number of explicit brick requests to the robot;
- number of explicit color suggestion requests to the robot;
- number of errors feedback from the robot to the user.

3.1.7 Subjective measures

Subjective measures were collected by asking participants to rate the degree to which they agreed or disagreed with each of 25 statements, using a 7-point Likert scale [strongly disagree (1), disagree (2), somewhat disagree (3), neutral (4), somewhat agree (5), agree (6) and strongly agree (7)]. The statements, reported in Table 1 were partially inspired by those used in previous studies (e.g.17-19), were aimed at assessing how participants: (i) perceived the task in terms of difficulty, physical and mental demands, time pressure (evaluating the statement 1 to 4); (ii) felt while performing the task (statements 5-9); (iii) perceived the robot in terms of appearance, behavior, trust, usefulness, and quality of the interaction (statements 10-25).

Following the passive condition, participants were administered only the first 9 statements, plus a session specific statement to assess whether they would have preferred to receive help during the task execution. Instead, the same 25 questions were administered at the end of both the reactive and proactive conditions. After each session, together with the questionnaire, each participant was also asked to estimate the number of errors they thought they committed. At the end of the whole experiment, participants had to indicate which of the experienced robot level of collaboration they preferred and which they considered the most useful.

3.2 Results

All statistical analyses were performed using IBM SPSS Statistics (Version 27).

3.2.1 Performance

The time to complete the task was analyzed by means of a repeated-measures Analysis of Variance (ANOVA) with Condition as within-participant factor and Order of presentation of the condition as between-participant factor. The alpha value (a priori criterion for the probability of falsely rejecting your null hypothesis) was set to 0.05.

The analysis showed only a main effect of Condition, $F(2,6) = 38.60$ and $p < 0.001$, with a faster execution time in the passive condition as compared to the other two. No significant difference emerged between the reactive and proactive conditions ($p = 0.47$).

Participants showed a general inclination towards using the robot. In the reactive condition, 88% of the participants asked the robot to bring a LEGO piece more than 5 times, one participant asked only 1 time, and 1 participant never asked. Likewise, in the same condition, 50% of the participants asked suggestions about the color at least 1 time, 3 participants asked twice, and 1 participant asked 4 times. In the proactive condition, only 1 person asked for help. Indeed, the participants did not seek help directly but relied on the automatic error feedback of the robot that was provided in 22% of the proactive sessions with a maximum of 3 times during a single execution. When estimating their accuracy level in performing the task, participants overestimated their error rate in the passive condition (12 predicted vs. 6 committed error) that was also perceived as more difficult, as we will discuss in the following section, while they underestimated it in the reactive modality (6 predicted vs. 10 committed errors).

3.2.2. Subjective measures

The mean responses to the Likert scale for the three experimental conditions are reported in Table 1.

Statements		Execution Modalities		
#	Text	PA	R	PR
1	The task was easy to accomplish	6.1 ± 1.4	6.3 ± 0.8	6.8 ± 0.5
2	The task was mentally demanding	3.1 ± 2.0	2.8 ± 1.8	2.4 ± 1.7
3	The task was physically demanding	3.2 ± 2.0	1.7 ± 1.4	1.6 ± 1.2
4	I felt pressure in accomplishing the task	3.1 ± 2.0	2.9 ± 2.0	2.7 ± 2.0
5	I felt in control of the outcome	5.8 ± 1.7	6.2 ± 1.1	6.0 ± 1.7
6	During the task, I felt anxious	2.5 ± 1.8	2.5 ± 1.8	1.9 ± 1.1
7	During the task, I felt comfortable	5.9 ± 1.5	6.1 ± 1.4	6.2 ± 1.2
8	During the task, I felt bored	2.4 ± 1.6	2.0 ± 1.5	2.3 ± 1.6
9	I was accurate in performing the task	6.1 ± 1.4	6.4 ± 0.9	6.7 ± 0.6
10a	I would have preferred to receive some help	5.3 ± 2.0	-	-
10b	The robot was helpful in accomplishing the task	-	6.4 ± 0.7	6.3 ± 1.3
11	The robot and I equally contributed to the task	-	4.5 ± 2.1	4.8 ± 2.2
12	The robot contribution to the task was higher	-	3.2 ± 2.2	3.1 ± 2.2
13	The interaction with the robot felt natural	-	5.8 ± 1.5	6.4 ± 0.9
14	The robot promptly reacted to my requests	-	6.1 ± 1.4	6.2 ± 1.1
15	The robot was accurate in perceiving	-	6.1 ± 1.4	6.7 ± 0.5
16	The robot's behavior met the needs of the task	-	5.9 ± 1.6	6.4 ± 1.3
17	The robot was too slow	-	2.6 ± 1.7	2.9 ± 2.0
18	I felt in competition with the robot	-	1.2 ± 0.6	1.3 ± 0.5
19	I felt the robot was reliable	-	5.7 ± 1.6	6.2 ± 1.0
20	I trusted the robot	-	6.2 ± 1.2	6.5 ± 0.6
21	The robot's behavior was easy to understand	-	6.4 ± 1.2	6.7 ± 0.6
22	I liked the robot's appearance	-	6.1 ± 1.1	6.0 ± 1.2
23	I felt comfortable in interacting with the robot	-	6.1 ± 1.3	6.6 ± 0.5
24	I felt safe in interacting with the robot	-	6.4 ± 0.9	6.7 ± 0.5
25	I am willing to interact again with the robot	-	6.2 ± 1.4	6.2 ± 1.4

Table 1: Mean responses to the Likert scale and standard deviation for the three experimental conditions.

To compare the 3 interaction modalities, numerical responses were submitted to two-tailed paired t-tests with the alpha value set at 0.05.

Comparisons with the PA condition could only be performed for the first 9 statements, since in this condition participants did not interact with the robot. Results showed a significant difference between the PA and the PR conditions in statement 1 about the task difficulty, $t(17) = -3.198$, $p = 0.005$. Another difference worth noticing can be appreciated between the PA and both R and PR conditions in statement 3, regarding the physical demands, $t(17) > 3.218, p < 0.005$ showing substantially different perceived physical effort needed for task completion. Concerning the last question in the PA condition, about the user indication on whether it would have liked to have received some help during the task, 61% of participants (i.e. 11 out of 18) selected the scores 6 (agree) and 7 (strongly agree).

Only two participants selected the score 1 (strongly disagree). When comparing the R and PR conditions, significant differences were again found for statement $1, t(17) = -2.297,\ p = 0.035$, with a higher score in the PR (6.8) than in the R (6.3) condition, 6, indicative of a decrease in the perceived task difficulty going from a reactive to a proactive robot behavior. Further marginally significant differences were found for statement 13 about the naturalness of the interaction, $t(17) = -1.975, p = 0.07$; and statement 15 about the robot accuracy in perceiving the user needs, $t(17) = -2.021,\ p = 0.06$; with higher scores in the PR conditions in both cases. The perceived trust in the robot and safety of collaboration did not significantly change between R and PR conditions. These questions are arguably tightly linked to the robot's single interacting actions and specific motion patterns, which did not change across the two conditions. Along this line, it must be noted that, although the robot did not display any error in its logic thanks to the Wizard-of-Oz paradigm, occasional errors occurred (less than 5% of the interactions) when grasping the blocks. These have been simply managed by swiftly attempting to grasp again limiting any possible negative effect on the perceived reliability of the robot.

In the questions administered at the end of the entire experiment, when asked about the preferred condition, 67% of participants (12 out of 18) indicated the PR as the most preferred condition, with 28% of the participants (5 out of 18) selecting the R and only one the PA condition. Similarly, the PR condition was indicated as the most useful by 78% of participants (14 out of the 18) with the remaining part indicating the R condition as the most useful robot behavior. This last choice was motivated by 3 out of 4 participants who felt to be "more in control" of the robot's actions and able to optimize it to their approach to the task in the R condition.

3.2.3. Summary of the main results

The present user study was conducted to address three research questions:

1. What is the level of collaboration that humans prefer in an HRI collaborative task?

2. What is the impact of the robot's behavior on the human perception of the task and robot?

3. Do humans wish to collaborate with a robot trading longer executions for a better experience?

As regards the first question, we found that out of 18 users, 17 of them preferred to use the robot's help; among these, 12 reported to prefer to interact with a proactive robot, while 5 reported to prefer to interact with a reactive robot.

As regards the second question, compared to the passive condition in which users performed the task by their own and no interaction with the robot occurred, the reactive robot behavior proved to be effective in decreasing the perceived difficulty and effort needed to complete the task, as witnessed by the statistically significant differences in subjective scores. Importantly, proactive behavior produced even more beneficial effects.

Finally, as regards the last question, our results clearly indicate that users preferred to receive help from the robot even though this lengthened the time taken to complete the task.

5. Conclusions

Some other general conclusions can be drawn from the results of the user study. Although the time to complete the task was significantly higher in the R and PR conditions, their task scores show greater values compared to the PA condition. This indicates that users were inclined to trade speed of execution for a generally more comfortable experience. Likewise, the robot score values were on average higher for the PR condition as compared to the R condition, which indicates that properly designed proactive behavior can provide the user a more positive experience during the interaction.

It should be noted that some participants indicated the R condition as both the preferred and most useful one, suggesting that a single approach may not fit all users. Nonetheless, the subjective evaluations provided by the participants were positive overall, and did not significantly change across modalities, except for two statements. Participants did not experience negative feelings while interacting with the robot with all but one participant expressing the will to interact again with the robot. Interestingly, participants felt in control of the outcome in both the R and PR conditions. This result suggests that well-calibrated proactive behavior could also help in minimizing the risk of users feeling not in control of the situation.

In the future, UNIMORE and SUPSI plan to validate whether similar results can be found when users are engaged in more complex tasks involving, for instance, robot navigation in cluttered environments, one of the requirements of SERMAS agents. Furthermore, we will also perform the same user study considering platforms with different sizes and types, to evaluate the role played by the robot’s appearance in the users’ acceptance. Finally, we will include in the evaluation metrics information about the user’s pose, facial expression, and gaze so to automatically draw conclusions about the users’ emotions. Results of these investigations would provide important indications of the characteristics the agents developed in the SERMAS project should possess to assure their social acceptance by users.

6. References

[1] Gaudiello I, Zibetti E, Lefort S et al. (2016) Trust as an indicator of robot functional and social acceptance. An experimental study on user conformation to Icub answers. Computers in Human Behavior 61: 633-655

[2] OpenAI (2023) Openai GPT-4 documen- tation. URL https://www.openai.com/research/gpt-4, accessed on 2023

[3] Ye Y, You H, Du J (2023) Improved trust in human-robot collaboration with ChatGPT. IEEE Access (2023).

[4] Ivanov S (2019) Ultimate transformation: how will automation technologies disrupt the travel, tourism and hospitality industries? Zeitschrift für Tourismuswissenschaft 11(1):25–43

[5] Naneva S, Sarda Gou M, Webb TL, et al (2020) A systematic review of attitudes, anxiety, acceptance, and trust towards social robots. International Journal of Social Robotics 12(6):1179–1201

[6] Song B, Zhang M, Wu P (2022) Driven by technology or sociality? use intention of service robots in hospitality from the human–robot interaction perspective. Inter- national Journal of Hospitality Management 106:103278

[7] Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly: 319-340

[8] Wirtz J, Patterson PG, Kunz WH, et al (2018) Brave new world: service robots in the frontline. Journal of Service Management 29(5):907–931

[9] Breckler SJ (1984) Empirical validation of affect, behavior, and cognition as distinct components of attitude. Journal of Personality and Social Psychology 47(6):1191

[10] Rossi A, Dautenhahn K, Koay KL, et al (2020) How social robots influence people's trust in critical situations. In: IEEE Int. Conf. on Robot and Human Interactive Communication, pp 1020–1025

[11] Kraus J, Miller L, Klumpp M, et al (2023) On the role of beliefs and trust for the intention to use service robots: An integrated trustworthiness beliefs model for robot acceptance. International Journal of Social Robotics pp 1–24

[12] Sirithunge C, Jayasekara ABP, Chandima D (2019) Proactive robots with the percep- tion of nonverbal human behavior: A review. IEEE Access 7:77308–77327

[13] Abbate G, Giusti A, Schmuck V, et al (2024) Self-supervised prediction of the intention to interact with a service robot. Robotics and Autonomous Systems 171:104568

[14] Yousuf MA, Kobayashi Y, Kuno Y, et al (2013) How to move towards visitors: A

model for museum guide robots to initiate conversation. In: 2013 IEEE RO-MAN, IEEE, pp 587–592

[15] Mead R, Matarić MJ (2015) Proxemics and performance: Subjective human evaluations of autonomous sociable robot distance and social signal understanding. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, pp 5984–5991

[16] Liu P, Glas DF, Kanda T, et al (2018) Learning proactive behavior for interactive social robots. Autonomous Robots 42:1067–108

[17] Akalin N, Kristoffersson A, Loutfi A (2022) Do you feel safe with your robot? factors influencing perceived safety in human-robot interaction based on subjective and objective measures. International Journal of Human-Computer Studies 158:102744

[18] Baraglia J, Cakmak M, Nagai Y, Rao RP, Asada M (2017). Efficient human-robot collaboration: When should a robot take initiative? International Journal of Robotics Research 36: 563-579

[19] Hoffman G (2019) Evaluating fluency in human–robot collaboration. IEEE Transactions on Human-Machine Systems 49(3):209– 218