Effects of Robot's Adaptive Autonomy on Users Experience in a Museum Scenario

Filippo Cantucci[1,*], Marco Marini[2] and Rino Falcone[3]

[1]Institute of Cognitive Sciences and Technologies, National Research Council of Italy (ISTC-CNR), 00185 Rome
[2]IMT School for Advanced Studies Lucca, 55100 Lucca, Italy
[3]Institute of Cognitive Sciences and Technologies, National Research Council of Italy (ISTC-CNR), 00185 Rome

Abstract

Integration of robots in society implies the need to design systems with adaptive behavior, especially in social domains like cultural heritage, healthcare, and education. This study examines user satisfaction when interacting with a robot guided by a computational cognitive model that incorporates Adjustable Social Autonomy principles. We conducted a within-subjects experiment in the Cultural Heritage domain, where museum visitors interacted with the humanoid robot Nao. The robot's task was to recommend a museum exhibition to visit, exercising different degrees of autonomy based on its capability to have a theory of mind. The findings showed that as the robot's autonomy in task adoption increased, user satisfaction with the robot decreased, while satisfaction with the tour itself improved. These results underscore the potential of adjustable social autonomy for developing autonomous adaptive social robots that enhance user experiences in Cultural Heritage scenarios like museums.

Keywords

Human-Robot Interaction, Autonomy Adaptation, User Experience, Cognitive systems, Theory of Mind

1. Introduction

To foster the acceptance of robots in society, it is crucial to deploy systems capable of adapting their behavior both to the environment and the needs of interacting users [1]. This is challenging across multiple social domains, including cultural heritage [2, 3], elderly assistance [4], and tourism [5, 6], among others. Despite significant advancements in adaptivity within Human–Robot Interaction (HRI) [7] and Social Robotics [8], research overlooked the critical role of analyzing the robot's cognitive processes and their implications for intelligent interaction, such as adaptive behavior [7], personalization, trust [9, 10] and so on. Therefore, it is essential to explore alternative interaction methodologies that facilitate the design of robots aligning with human expectations, capabilities, and limitations within specific contexts. The starting point is to identify the priority needs that humans expect to satisfy during effective interactions with their peers and, consequently, to shift our focus to HRI scenarios. Humans need to i) pursue their own goals, i.e., the state of affairs the user wants to achieve; ii) consider broader interests or active goals, i.e., the state of affairs that the user, with a specific profile and set of mental

WOA 2024: 25th Workshop "From Objects to Agents", July 8-10, 2024, Forte di Bard (AO), Italy

*Corresponding author.
filippo.cantucci@istc.cnr.it (F. Cantucci); marco.marini@imtlucca.it (M. Marini); rino.falcone@istc.cnr.it (R. Falcone)
0000-0003-2972-9666 (F. Cantucci)

states, could be interested in or has already planned to achieve. The first consideration implies achieving goals in the context of the ongoing interaction. In HRI, this determines building robots capable of understanding these goals and selecting an appropriate strategy (e.g., plan) to accomplish them. The level of user satisfaction will depend on the plan the robot intends to execute and the boundary conditions it employs. The second consideration arises whenever humans have interests and goals that extend beyond the current interaction. To be an effective collaborator, a robot should comprehend these implicit interests and goals and work towards their achievement or protection. For instance, a task delegated by a human could be part of a much more complex goal, or the same task might conflict with other goals that the human has not consciously considered. In both scenarios, the robot's role goes beyond simple task execution and requires a more complex assessment of the human's mental states (e.g., intentions, beliefs, motivations, interests); the robot must select a plan that aligns with those mental states. This can imply a mismatch between the task delegated and the one adopted [11].

The present study investigates human user satisfaction when interacting with a robot whose decision-making process is guided by a computational cognitive model [2] integrating the principles of adjustable social autonomy [12]. We designed a within-subjects experiment in cultural heritage, where museum visitors interacted with the humanoid robot Nao [13]. The robot's task was to provide the user with a museum exhibition to visit. The rest of the paper is organized as follows: Section 2 provides a theoretical background on the approach exploited in the present work; Section 2.1 gives an overview of the state of the art in adaptivity in HRI and in the specific domain of museums; Section 3 presents the experimental methods applied in our experiment with the Nao robot; in Section 4, we describe the statystical analysis exploited for analyzing the results; furthermore we presents the experimental findings; in Section 5 we discuss the results obtained; Section 6 is dedicated to conclusions and future work.

2. Background and Related Works

The theoretical models mentioned in the introduction have been applied to create a cognitive computational model [14, 2], defining the behavior of the robot in this experiment. Without going into detail, the proposed cognitive model defines a task-oriented, belief–desire–intention (BDI) [15, 16] robot that performs effective reasoning within dynamic interaction with other cognitive agents, typically humans. The robot adopts a delegated task τ and exerts some degree of discretion, which corresponds to various levels of autonomy in τ adoption. The capability to expand the decision and action space concerning τ shows the adaptability to the user's mental states as a criterion. In fact, this adaptability is possible thanks to the robot's capability to have a ToM, enabling it to attribute mental states beyond those explicitly declared (we defined them as active goals). The user's model is built following a human profiling phase, during which the robot extracts relevant user features via verbal and non-verbal communication. After that, the robot elicits an internal negotiation process in which it mediatesτ adoption considering constraints imposed by the ongoing interaction context (e.g., level of crowding in a room, other agents' mental states). Finally, the cognitive model gives the robot the capability to explain the reasons that led it to adopt τ at a specific level of autonomy.

According to this cognitive model and Adjustable Social Autonomy theory, the robot possesses the capacity to adopt delegated tasks at various autonomy levels. For instance, the robot can execute the exact task delegated by the user, providing literal help. Otherwise, it can modify the user's plan, considering additional user interests or active goals. This results in the adoption of a more complex form of assistance, such as critical help or over help. Sub-help, on the other hand, occurs every time the robot fails to accomplish the task adopted in literal help, but still manages to achieve a sub-goal related to the task.

2.1. Adaptive Social Robots in Museum Settings

Cultural heritage sites such as museums represent a complex scenario, suitable for robots that are able to assist and interact with people in a natural manner. Different pioneering work [17, 18, 19, 20, 21] has been proposed, with the goal of designing robots able to be deployed in a museum and to integrate different HRI capabilities. Despite their pioneering and remarkable work in autonomous social navigation and interaction with humans, these approaches do not realize a real behavior adaptation [22] to the user experience. Multiple works [23] proposed technical methods to integrate social navigation, user perception, and verbal and non-verbal communication in order to adopt different behaviors on the basis of the users involved in the interaction. Recent works [24, 25, 26] tried to implement a much more effective and personalized user experience by designing robots that are able to exploit their perceptive and decision-making skills to establish much deeper interactions with visitors. Our approach is to try to achieve a real user's adaptation, where the robot engages in reasoning based on the task delegated by the user. This reasoning process is grounded in complex models of both the interacting user (theory of mind) and the interaction mode, as described in the introduction of this work.

3. Methodology

To assess the impact of the adjustable social autonomy paradigm on HRI, we conducted an experiment in a real museum. The present experiment involved interactions between human users (museum visitors) and the humanoid robot Nao, acting as a museum assistant. During the experiment, Nao provided museum visitors the option to embark on a virtual museum tour through a computer monitor at the museum's exit. The user-assigned task to the robot was to provide assistance by offering a guided museum tour that showcased artworks from an explicitly preferred artistic period. Notably, the artworks featured in the virtual tour were unrelated to those in the actual museum exhibition. Nao's primary goal was to assist visitors in selecting the museum exhibition to visit. Subsequently, Nao guided users through the selected tour. The present experiment focused on analyzing the impact of intelligent help on visitor satisfaction and the level of surprise as indicators of the robot's ability to intercept the visitors' needs, even when not explicitly declared. In particular, we investigated the visitors' satisfaction with (1) the suggested tour (adoption results) and (2) the robot's behavior (adoption strategy).

As mentioned earlier, the robot employed different levels of task adoption in order to offer optimal support to the user. Specifically, two types of help were exploited in this experiment: literal help and critical help. When the robot opted for literal help, it constructed the tour by selecting the most relevant artworks from the artistic period explicitly indicated by the user. Conversely, with critical help, the tour was crafted based on broader criteria, considering additional user profile information such as tolerance for room crowding or disinterest in specific artistic periods. The robot considered the potential user's interest in highly relevant artworks within the virtual museum, not necessarily belonging to the artistic period specified in the task delegation; this aspect was implicitly inferred by the robot and not explicitly stated by the user. Specifically, the robot aimed to optimize the correlation between the relevance of the artworks and the simulated room crowding where they were placed. The strength of this form of assistance lies precisely in the robot's ability to go beyond the goals declared by the user and address other needs and interests that the user may not immediately consider. However, it is crucial to note that this type of assistance is susceptible to potentially erroneous robot interpretations and could be met with reluctance from the user regarding a tutoring role played by the robot, which was not directly requested. We aimed to verify the following research hypotheses:

- $\boldsymbol{H_1}$: user satisfaction regarding the quality of the tour suggested by the robot was higher when the robot provided critical help, as opposed to literal help;
- $\boldsymbol{H_2}$: users exhibited greater satisfaction with the robot's behavior when the robot operated in critical help, rather than providing literal help;
- $\boldsymbol{H_3}$: users experienced a higher level of surprise with the robot's selection when it performed critical help compared to literal help

Although $\boldsymbol{H_1}$ specifically pertains to satisfaction regarding the robot's performance (suggested tour), $\boldsymbol{H_2}$ aims to investigate participant satisfaction with the robot's behavior. This behavior is intended as the decision-making process that led to the adoption of the task by performing either literal help or critical help. The participants precisely evaluated how the robot operated concerning the task initially delegated to it, regardless of the satisfaction linked to the outcome obtained following its behavior.

We recruited 84 participants for the purpose of this study. In order to reduce ecological concerns, the experiment was conducted in an authentic museum setting, namely the Palazzo delle Esposizioni [27], located in Rome. The user sample consisted of 49 men and 35 women. A total of 28 of the participants ranged between 18 and 35 years old, 54 between 36 and 65 years old, and 2 were older than 65 years old. Participants reported their level of expertise in the artistic domain using a 5-item Likert scale ranging from lower expertise to maximum expertise (M = 2.655, SD = 0.857). Prior to their participation in this study, each participant approved an informed consent. We conducted a human-participant experimental study in a within-subject fashion. We implemented two experimental conditions in a counterbalanced order: in the control condition (LH), the robot provided literal help to the user. During the experimental condition (CH) the robot performed critical help (see Section 3 for literal help/critical help definitions).

Figure 1 illustrates the scenario's design, which was specifically created in a dedicated room of the Palazzo delle Esposizioni. After visiting the exhibitions available in the building, each user had free access to this section in order to participate in the experiment. As the figure shows, the robot is in front of the user, near a screen displaying a web interface. Through this interface, the user can respond to questions about her artistic preferences posed by the robot and explore the virtual tour (Figure 2). The virtual tour consists of a selection of artworks, selected by the robot, according to the criteria exposed in Section 3. In addition to guiding the user through various phases of the experiment, the robot describes (via speech) each artwork,

Figure 1: Experimental scenario representation. Not an actual participant.

Figure 2: The figure shows an example of how an artwork is exposed to the user. (Image credits: Galleria Borghese/photo Luciano Romano).

providing different levels of detail depending on the user's initial preferences. During the visit, the user may encounter noisy conditions, similar to those experienced during a real museum visit. For instance, artworks may be corrupted by simulated noise to mimic the presence of other users viewing the same artwork (Figure 3). Additionally, the robot's descriptions of the artworks may be superimposed on background noise of varying intensity, simulating the

Figure 3: Noise simulation in the artwork. (Image credits: Galleria Borghese/photo Luciano Romano).

crowded environment of the room where the user is viewing the artwork.

The investigation of user satisfaction and their level of surprise involved the following experimental phases:

1. *Starting Interaction*: the robot started the interaction by introducing itself to the user, providing an overview of its role and the virtual museum it managed;
2. *User Profiling*: The robot investigated the user's artistic preferences in terms of favored and less favored artistic periods. Subsequently, the robot asked the user about the desired level of detail in the artwork descriptions as an indicator of the user's intended level of engagement. Lastly, the robot asked the user to specify her tolerance level regarding the number of people present in a room while viewing an artwork (crowding index).
3. *First Tour presentation*: After finalizing the user's profile, the tour composition was determined by the condition the robot randomly decided to implement (LH or CH). Once the selection was complete, the robot activated the corresponding tour and left the control to the user, allowing her the virtual tour.
4. *First Tour evaluation*: at the conclusion of the first virtual tour, the robot administrated a questionnaire to assess the user satisfaction and surprise level with the completed visit.
5. *Second Tour presentation*: subsequently, the robot offered an alternative tour, implementing the opposite type of help compared to the one randomly selected in the first tour.
6. *Second Tour evaluation*: after the completion of the second tour, the robot administrates the user with a new satisfaction/surprise questionnaire.

For the robot's decision-making system, we utilized an agent-oriented programming (AOP) software known as Jason [28, 29], a widely adopted tool for programming artificial agents. The computational model underlying the robot's decision-making process is detailed in [2]. To develop the web application (i.e., virtual museum), we opted for the Java-based Spring Boot framework [30]. Furthermore, the robot employed in this study is the humanoid Nao robot [13], extensively used in HRI scenarios. The Nao robot operates on a specialized Linux-based operating system known as NAOqi, which powers the robot's multimedia capabilities. These capabilities include four microphones for voice recognition and sound localization, two speakers for multilingual text-to-speech synthesis, and two HD cameras for computer vision tasks such as facial and shape recognition. We collected a mySQL database of 344 artworks, organized in artistic periods in such a way that it covers the entire body of the history of art. The categorization of the history of art periods is based on the work of one of the most important art historians of the 20th century, Giulio Carlo Argan [31]. Artworks are evenly distributed among all artistic periods. All the employed images had a creative commons license and could be uses for any purpose.

Participants responded to a questionnaire administered at the conclusion of both control and experimental conditions. Through this questionnaire, we aimed to measure the following variables:

1. Tour Satisfaction: User satisfaction with the presented tour was measured using the following question: "How satisfied are you with the quality of the visit that the robot offered you?" Participants responded using a 5-item Likert scale ranging from not satisfied at all (1) to completely satisfied (5);
2. *Robot Behavior Satisfaction*: For assessing user satisfaction with the robot's behavior, participants were asked the following question: "How satisfied are you with the choice of robot compared to the artistic habits you expressed at the beginning of the interaction?" Participant satisfaction was investigated the same as with Tour Satisfaction;
3. *Surprise*: For assessing the users' surprise about the type of help provided by the robot, we exploited the question "How surprised are you by the choices made by the robot?" Participants responded using an 11-item Likert scale. A rating of 5 indicated a strong negative surprise, 0 represented no surprise at all, and 5 signified complete positive surprise.

4. Results

The statistical analyses were carried out with IBM SPSS 27.0 [32] and R through Jamovi [33]. A significance level of $\alpha = 0.05$ was selected for all analyses and p-values were reported. When significant effects were identified, the results include the unstandardized estimate (β), the 95% Confidence Intervals (CI), and the standard error of the estimate (SE). Before running the analyses, we assessed the normality of the model-dependent variables (i.e., Tour Satisfaction) using the Shapiro–Wilk test. The results highlighted statistical significance (all p-values < 0.001), suggesting a substantial deviation from normality. More specifically, both Tour Satisfaction and Robot Performance Satisfaction were negatively skewed (Skewness = -0.24; -0.09). For these reasons, we employed Generalized Mixed Effect Models [34]. In the case of skewed positive values, GLMM analysis was executed using the GAMLj module [35] using a Gamma distribution with a log link function [36]. These models reported the best Akaike's and Bayesian Information criteria (AIC and BIC) as compared with GLMM built using identity and inverse link functions. Lastly, all the models included participants' intercepts and slopes as random effects. This approach accounted for the inherent variations wherein participants could exhibit diverse

Table 1
Descriptive statistics of the main dependent variables.

	Literal Help			Critical Help		
	N	Mean	SD	N	Mean	SD
Tour Satisfaction	84	3.17	1.00	84	3.52	1.11
Robot Performance Satisfaction	84	3.82	1.04	84	3.48	0.94
Surprise	84	3.25	0.83	84	3.49	0.88

baseline levels of Satisfaction and could show singular changes in Satisfaction as a function of the Help Type. The main descriptive statistics are summarized in Table 1.

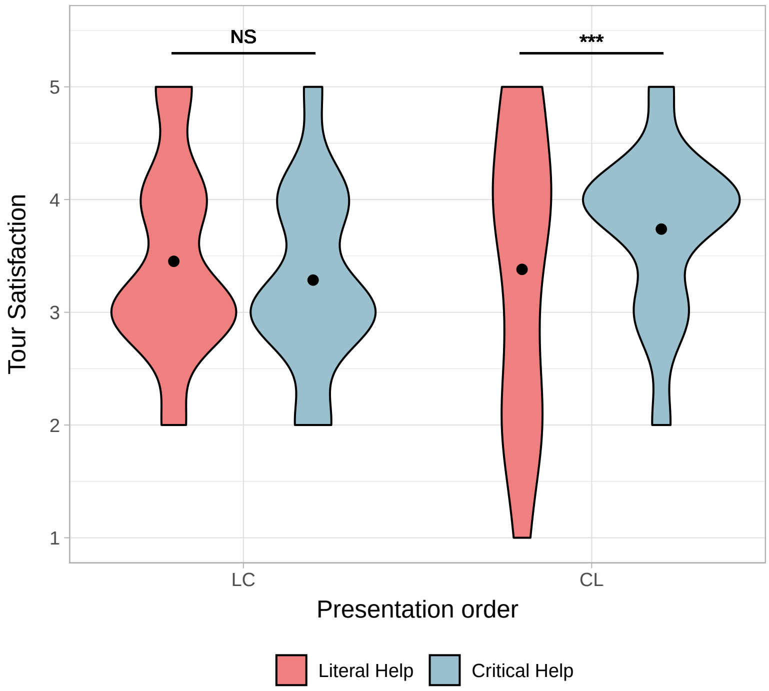

Figure 4: Tour Satisfaction as a function of Help Type and Presentation Order. *** $p < 0.001$; NS, Not Significant.

To assess the influence of Help Type (literal help or critical help) on users' self-reported Tour Satisfaction ($\boldsymbol{H_1}$), we conducted a Generalized Linear Mixed Model (GLMM) with a Gamma distribution and a logarithmic link function (see Figure 4). Help Type and Presentation Order (i.e., literal help first or critical help first) were modeled as fixed factors, whereas the Robot Behavior Satisfaction score was included as a covariate. Additionally, we incorporated individual variability by including participants' intercepts and slopes as ran-

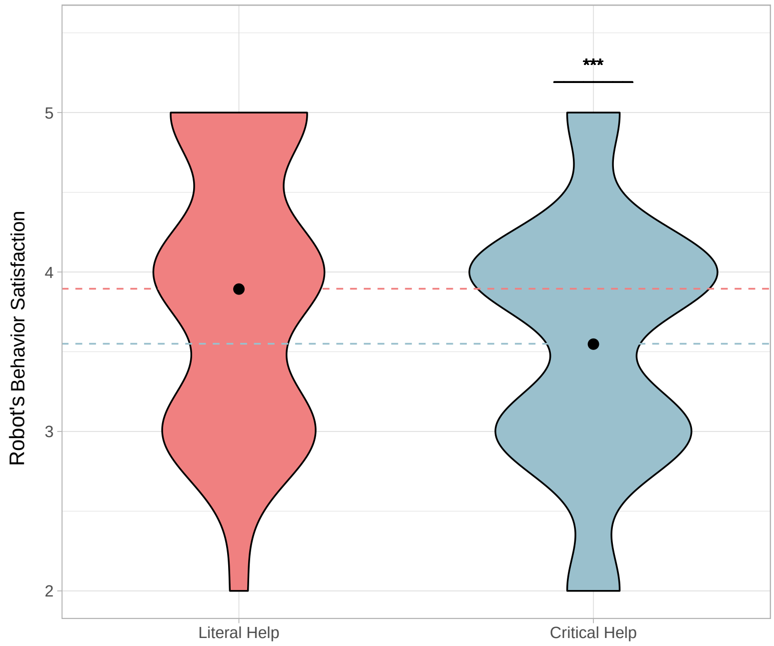

Figure 5: Robot Behavior Satisfaction as a function of Help Type. The violin plot visualizes the distribution curve. The horizontal dashed lines represent the median values. *** p < 0.001.

dom factors. In our model, Help Type exhibited a positive impact ($\beta = 0.11, SE = 0.03$, $95\% CI[0.05, 0.17], p < 0.001$) on Tour Satisfaction, indicating that a specific robot strategy (critical help) positively influenced satisfaction levels. Similarly, Robot Behavior Satisfaction had a significant positive effect ($\beta = 0.16, SE = 0.02, 95\% CI[0.12, 0.21], p < 0.001$), suggesting that greater satisfaction with the robot's behavior during the visit predicted a higher Tour Satisfaction. In contrast, the main effect of Presentation Order was not significant ($p = 0.91$). However, we detected a significant interaction effect between Help Type and Presentation Order ($\beta = 0.17, SE = 0.05, 95\% CI[0.06, 0.03], p = 0.003$), indicating that the impact of Help Type on Tour Satisfaction depended on the order in which it was presented during the visit experience. Subsequent Bonferroni-corrected post hoc analyses revealed a statistically significant increase in Tour Satisfaction exclusively among participants who received critical help followed by literal help ($p < 0.001$), compared to those who received literal help followed by critical help ($p = 1.00$) (Figure 4).

Secondly, to investigate the impact of Help Type on Robot Behavior Satisfaction ($\boldsymbol{H_2}$), we conducted a GLMM using a Gamma distribution and a logarithmic link function (see Figure 5). This analysis also included Presentation Order and the Help Type × Presentation Order interaction as fixed factors. The analysis revealed a significant effect for Help Type ($\beta = -0.09, SE = 0.03, 95\% CI[-0.15, -0.04], p < 0.001$), indicating that the critical Help Type had a negative effect on Robot Behavior Satisfaction compared to the literal help type. Conversely, both the main effect of Presentation Order and the interaction effect were not statistically significant ($p = 0.149$; $p = 0.242$).

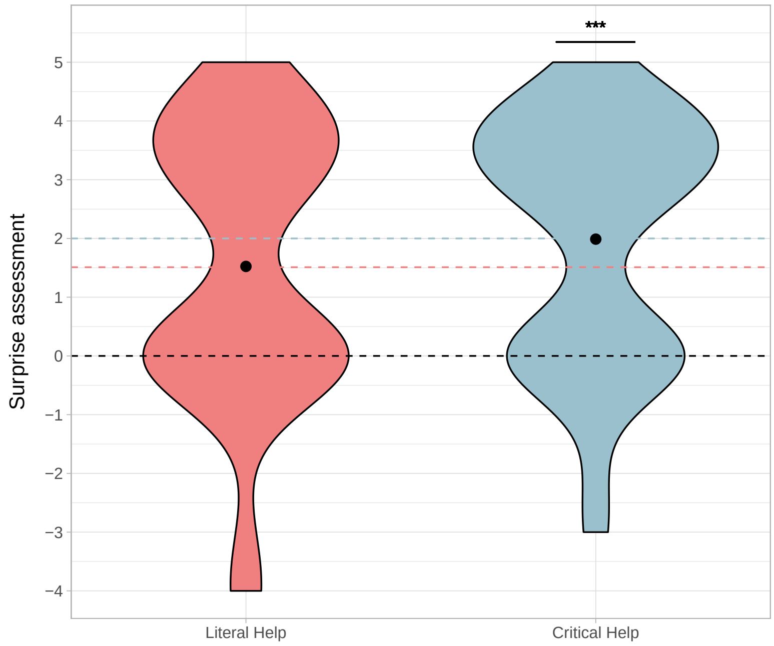

Figure 6: Surprise as a function of Help Type. The violin plot visualizes the distribution curve. The horizontal dashed lines represent the median values. *** p < 0.001.

Finally, to examine the surprise levels after receiving different types of help ($\boldsymbol{H_3}$), we performed a Generalized Linear Mixed Model (GLMM) (see Figure 6). In this analysis, Help Type and Presentation Order were treated as fixed factors, and Robot's Performance Satisfaction was included as a covariate. As in the previous models, participants' intercepts and slopes were considered random factors. Firstly, Robot Performance Satisfaction, representing the perceived satisfaction for the robot's suggestions during the visit, exhibited a significant positive effect ($\beta = 0.88, SE = 0.20, 95\% CI[0.49, 1.26], p < 0.001$) on the participants' surprise. This finding suggests that greater satisfaction with the robot's performance was associated with higher levels of surprise. Additionally, Help Type reported a significant positive effect ($\beta = 0.76, SE = 0.29, 95\% CI[0.21, 1.33], p = 0.009$), proving that participants who received critical help reported experiencing more surprise compared to those who received literal help. Furthermore, Presentation Order displayed a significant negative effect ($\beta = -0.92, SE = 0.36, 95\% CI[-1.63, -0.21], p = 0.013$). This suggests that the sequence in which the interaction occurred (between critical and literal help) influenced surprise levels, with lower levels of surprise observed when critical help preceded the literal one. Lastly, the interaction effect between Help Type and Presentation Order did not reach statistical significance ($p = 0.140$). In summary, this analysis unveiled distinct effects on participants' levels of surprise. Specifically, higher levels of Robot's Performance Satisfaction and receiving critical help were associated with increased levels of surprise, whereas receiving critical help at the beginning of the experiment reduced surprise scores.

5. Discussion

As exposed in the hypotheses defined in Section 3, our main research interest was to understand the effect of a robot employing its critical autonomous capacity in interactions with humans. In this paper, our investigation mainly focused on two key dimensions:

1. The effect on user satisfaction: we aimed to discern how the introduction of critical capacity affects user satisfaction ($\boldsymbol{H_1}$ and $\boldsymbol{H_2}$);
2. The level of surprise: we sought to determine the extent of surprise elicited in humans by this type of robot behavior ($\boldsymbol{H_3}$).

Our findings confirmed $\boldsymbol{H_1}$: user satisfaction with the quality of the proposed tour was greater when the robot developed a non-trivially literal solution, thereby introducing what we have defined as critical help, compared to a delegation aligned with the user's artistic preferences; namely, literal help. This finding was independent of the order in which the type of help was provided.

Hypothesis $\boldsymbol{H_2}$ focused on user satisfaction with the robot's performance in two types of task adoption: literal and critical help. Contrary to our expectations, the results revealed a preference towards the robot when it provided literal help rather than critical help, regardless of the order in which these levels of task adoption were introduced. It is essential to consider the following factors: (i) the average scores of user satisfaction for both literal and critical help were relatively high, even though the literal help received a higher rating; (ii) these results were not affected by the explanation that the robot provides after the interaction. The potential impact of these explanations will be further examined in a future study.

Regarding hypothesis $\boldsymbol{H_3}$, which concerned the level of user surprise in response to the robot's adoption strategies, results showed that users exhibited a relatively high level of surprise in both instances of literal and critical help. However, hypothesis $\boldsymbol{H_3}$ was confirmed as users showed greater surprise when they received critical help from the robot. The high degree of surprise even in the case of literal help likely stems from the intrinsic nature of the Human–Robot Interaction and the unexpectedly positive outcome arising from the robot's ability to select artworks based on the user's artistic preferences. This could also be explained by the fact that higher levels of satisfaction with the robot's behavior were associated with increased surprise.

Our results, taken as a whole, represent initial experimental evidence of the impact of the adjustable social autonomy model in HRI scenarios. The robot's capacity to attribute mental states to the user allows it to dynamically adapt the task delegated to the user's unexpressed needs. This is suggested by the fact that user satisfaction increases when they receive a type of museum tour that not only considers the artworks belonging to their preferred period, but also includes relevant works from other similar periods that were selected based on the user noise tolerance level. This represents an initial confirmation of the importance of introducing autonomous robots in terms of cognitive agents, namely agents possessing propositional attitudes, assuming an intentional stance, and having a representation of the other agent's mind. An autonomous robot that adapts its behavior may not be able to tailor it to a user's needs but may instead adopt other criteria, such as considering only the resources available in the physical world. This could lead the robot to make choices that potentially may have nothing to do with the user's expectations. This would result in user dissatisfaction with the outcome achieved by the robot. What we observed in our experiment is that a robot capable of adapting its behavior to the user's needs, including implicit mental states, can adopt a task in a way that does not decrease user satisfaction compared to the result obtained. Assistance that goes beyond literal interpretation satisfies the user more than a kind of help that merely considers what the user has explicitly expressed. These data represent further confirmation of the adjustable social autonomy model: robots that adopt tasks at different levels of autonomy change their role from being systems capable of executing even complex tasks to becoming collaborators (i.e., agents who take the initiative to provide unexpected behavior that diverge from the delegated task, either by proposing or directly executing different actions.

A higher satisfaction with the robot's behavior when it provides literal help can be attributed to the inherent risks associated with assistance that goes beyond explicit delegation, as it can generate misunderstandings, doubts, susceptibility, and refusals. This observation is corroborated by the finding that users are much more satisfied when the robot provides literal help. This confirms the potentially conflictual nature of collaboration when the robot adjusts its level of task adoption. This might also generate the perception of uncertainty regarding the endorsement of an autonomous robot's behavior. In this sense, it is relevant for the robot to evaluate the risks of conflicts that could arise due to inappropriate initiatives. Nevertheless, the tour satisfaction is higher in critical help than in literal help. In other words, the user exhibits higher satisfaction with the final result proposed by the robot when the robot adopts autonomy criteria in its collaborative efforts. However, the user may experience some degree of perplexity when assessing a robot that chooses to collaborate with a certain level of autonomy. Another confirmation of the positive yet equally conflicting impact of critical help lies in the fact that users are positively surprised by critical help.

6. Conclusions and Future Works

This paper describes how adjustable social autonomy can be used to systematically characterize the behavior of a robot when it adopts a task explicitly delegated by a human. The core of this approach is the robot's capability to attribute mental states to the delegating user and execute the task at various levels of autonomy. This is achieved by considering the user's mental states (goals, beliefs, expectations, and so on) that go beyond what is explicitly delegated. We tested the impact of this approach on user satisfaction and surprise, and although there are limitations to this experiment, the results are promising, demonstrating how this type of interaction can be both highly satisfying and surprising for the user. However, it is important to note that it can also lead to collaborative conflicts that may affect user satisfaction with the robot's behavior. These collaborative conflicts can be mitigated using various strategies. Results and their discussion point out the potential of exploiting adjustable social autonomy as a paradigm for developing computational cognitive models for adaptive social robots that can improve user experiences in multiple HRI real domains. Certainly, experiments of this nature should be conducted in various domains of social robotics to validate the robustness of this type of interaction across different application domains. This aspect will be one of the points that we will need to address in future work.

Acknowledgments

This work has been funded by the project FAIR - Future Artificial Intelligence Research (MIUR-PNRR).

References

[1] D. David, P. Thérouanne, I. Milhabet, The acceptability of social robots: A scoping review of the recent literature, Computers in Human Behavior 137 (2022) 107419.
[2] F. Cantucci, R. Falcone, Collaborative autonomy: Human–robot interaction to the test of intelligent help, Electronics 11 (2022) 3065.
[3] N. Gasteiger, M. Hellou, H. S. Ahn, Deploying social robots in museum settings: A quasi-systematic review exploring purpose and acceptability, International Journal of Advanced Robotic Systems 18 (2021) 17298814211066740.
[4] G. D'Onofrio, D. Sancarlo, M. Raciti, D. Reforgiato, A. Mangiacotti, A. Russo, F. Ricciardi, A. Vitanza, F. Cantucci, V. Presutti, et al., Mario project: Experimentation in the hospital setting, in: Ambient Assisted Living: Italian Forum 2017 8, Springer, 2019, pp. 289–303.
[5] F. Cantucci, R. Falcone, A computational model for cognitive human-robot interaction: An approach based on theory of delegation., in: WOA, 2019, pp. 127–133.
[6] S. Fang, X. Han, S. Chen, The impact of tourist–robot interaction on tourist engagement in the hospitality industry: A mixed-method study, Cornell Hospitality Quarterly 64 (2023) 246–266.
[7] M. I. Ahmad, O. Mubin, J. Orlando, A systematic review of adaptivity in human-robot interaction, Multimodal Technologies and Interaction 1 (2017) 14.
[8] H. Mahdi, S. A. Akgun, S. Saleh, K. Dautenhahn, A survey on the design and evolution of social robots—past, present and future, Robotics and Autonomous Systems 156 (2022) 104193.
[9] A. Sapienza, F. Cantucci, R. Falcone, Modeling interaction in human–machine systems: A trust and trustworthiness approach, Automation 3 (2022) 242–257.
[10] R. Falcone, A. Sapienza, F. Cantucci, C. Castelfranchi, To be trustworthy and to trust: The new frontier of intelligent systems, Handbook of Human-Machine Systems (2023) 213–223.
[11] F. Cantucci, R. Falcone, M. Marini, Redefining user expectations: The impact of adjustable social autonomy in human–robot interaction, Electronics 13 (2023) 127.
[12] R. Falcone, C. Castelfranchi, The human in the loop of a delegated agent: The theory of adjustable social autonomy, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 31 (2001) 406–418.
[13] A. Robaczewski, J. Bouchard, K. Bouchard, S. Gaboury, Socially assistive robots: The specific case of the nao, International Journal of Social Robotics 13 (2021) 795–831.
[14] F. Cantucci, R. Falcone, Towards trustworthiness and transparency in social human-robot interaction, in: 2020 IEEE International Conference on Human-Machine Systems (ICHMS), IEEE, 2020, pp. 1–6.

[15] A. S. Rao, M. P. Georgeff, et al., Bdi agents: from theory to practice., in: Icmas, volume 95, 1995, pp. 312–319.
[16] L. De Silva, F. R. Meneguzzi, B. Logan, Bdi agent architectures: A survey, in: Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI), 2020, Japão., 2020.
[17] W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, S. Thrun, Experiences with an interactive museum tour-guide robot, Artificial intelligence 114 (1999) 3–55.
[18] S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, J. Schulte, et al., Minerva: A second-generation museum tour-guide robot, in: Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C), volume 3, IEEE, 1999.
[19] M. Nieuwenhuisen, S. Behnke, Human-like interaction skills for the mobile communication robot robotinho, International Journal of Social Robotics 5 (2013) 549–561.
[20] A. Chella, M. Liotta, I. Macaluso, Cicerobot: a cognitive robot for interactive museum tours, Industrial Robot: An International Journal 34 (2007) 503–511.
[21] T. Willeke, C. Kunz, I. R. Nourbakhsh, The history of the mobot museum robot series: An evolutionary study., in: FLAIRS, volume 2001, 2001.
[22] M. K. Lee, J. Forlizzi, S. Kiesler, P. Rybski, J. Antanitis, S. Savetsila, Personalization in hri: A longitudinal field experiment, in: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, 2012, pp. 319–326.
[23] M. Hellou, J. Lim, N. Gasteiger, M. Jang, H. S. Ahn, Technical methods for social robots in museum settings: An overview of the literature, International Journal of Social Robotics 14 (2022) 1767–1786.
[24] J. Park, J. Kim, D.-Y. Kim, J. Kim, M.-G. Kim, J. Choi, W. Lee, User perception on personalized explanation by science museum docent robot, in: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), IEEE, 2022, pp. 973–975.
[25] T. Iio, S. Satake, T. Kanda, K. Hayashi, F. Ferreri, N. Hagita, Human-like guide robot that proactively explains exhibits, International Journal of Social Robotics 12 (2020) 549–566.
[26] F. Cantucci, R. Falcone, Autonomous critical help by a robotic assistant in the field of cultural heritage: A new challenge for evolving human-robot interaction, Multimodal Technologies and Interaction 6 (2022) 69.
[27] Palazzo delle esposizioni roma, https://www.palazzoesposizioni.it/, . Accessed on 1 May 2022.
[28] R. H. Bordini, J. F. Hübner, M. Wooldridge, Programming multi-agent systems in AgentSpeak using Jason, John Wiley & Sons, 2007.
[29] Jason, https://github.com/jason-lang/jason/releases, . Accessed: 2010-09-30.
[30] Springboot, https://spring.io/projects/spring-boot, . Accessed on 22 January 2022.
[31] G. C. Argan, Storia dell'arte italiana, (No Title) (1968).
[32] D. George, SPSS for windows step by step: A simple study guide and reference, 17.0 update, 10/e, Pearson Education India, 2011.
[33] Jamovi, https://www.jamovi.org/, . Accessed on 22 January 2022.
[34] W. W. Stroup, Generalized linear mixed models: modern concepts, methods and applications, CRC press, 2012.

[35] General analyses for the linear model in jamovi, https://gamlj.github.io/, . Accessed on 22 January 2022.
[36] V. K. Ng, R. A. Cribbie, Using the gamma generalized linear model for modeling continuous, skewed and heteroscedastic outcomes in psychology, Current Psychology 36 (2017) 225–235.