“God, grant me the serenity to accept the things I cannot change,

the courage to change the things I can,

and the wisdom to know the difference.”

- *Reinhold Niebuhr*

“To work and create 'for nothing', to sculpture in clay, to know that one's creation has no future, to see one's work destroyed in a day while being aware that fundamentally this has no more importance than building for centuries- this is the difficult wisdom that absurd thought sanctions. Performing these two tasks simultaneously, negating on one hand and magnifying on the other, is the way open to the absurd creator. He must give the void its colors.”

– *Albert Camus*

“Will you let me try once more?” – Social Avatars in Teleoperation to improve Trust in Automation

by

Pouya Pournasir

Bachelor of Science in Computer Engineering, University of Guilan, 2020

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Computer Science

in the Graduate Academic Unit of Computer Science

Supervisor: Daniel J. Rea, PhD, Faculty of Computer Science

Examining Board: Scott Bateman, PhD, Faculty of Computer Science
Eric Aubanel, PhD, Faculty of Computer Science
Veronica Whitford, PhD, Department of Psychology

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

December 2023

ABSTRACT

In robot teleoperation, operators and robots collaborate to achieve goals through shared autonomy, reducing operator workload. Trusting the robot is critical for the operator's success, as low trust results in operators not delegating tasks to the robot, even if doing the task themselves creates a high workload that reduces their performance. Research shows trust repair is challenging, but social HRI suggests that robots can rebuild trust in social situations through social strategies such as acknowledging mistakes and promising to do better. This study explores integrating these social strategies and interfaces into teleoperation to enhance trust repair. We compared a social cue-based interface to a conventional one, theorizing that adding in social cues would increase the effectiveness of social trust-repair strategies. Our study found that participants view the social cue-based interface as more capable, a factor that can make the participant put more trust on the system.

DEDICATION

I want to dedicate this thesis to my mother. Despite all the hardships and struggle, she tried her best to support me through my academic journey from school to university. I also want to dedicate this thesis to my father, whom in my worst times he encouraged me to move forward and helped me find the light in the dark. In the end, I will dedicate this thesis to my brother, as an ode to our memories together.

Without them, it was impossible to reach this stage of my life.

ACKNOWLEDGEMENTS

First and foremost, I want to thank my supervisor, Professor Daniel Rea. Words cannot express how deeply I am thankful for everything you have done. In each stage of my masters, from picking me up from airport when I arrived in Fredericton to this moment, he tried his best to support me in every aspect. His dedication to his research has always inspired me to strive for more and perform better. In each obstacle I have encountered in my research, he was there. Always ready to help and give feedback. It is difficult to not end your meeting with Daniel and not have a smile on your face. His positive attitude, sense of humor, and strong passion made our discussions the best part of my master's. I get never tired from listening to his opinion and experiences, even it is hours. I wish to maintain our new form of friendship in the future! True definition of a great mentor, a great father, and a great friend.

I owe a debt of gratitude to Parinaz, my gorgeous and compassionate partner. Parinaz, you were the only person, alongside Daniel, who have been in every stage of this journey with me. In my darkest times when it was difficult to maintain my sanity, when the sun gave her cold shoulder in the midst of the first winter, you were my source of light. You helped me to move forward, you brought order in utter chaos. We made a lot of memories, together, we laughed, we cried, we fought, we tried, we failed, and we learned, but in all those moments, I loved you. It's hard to imagine these two years without you, I wish you all the best.

I'm grateful for Kimchi, my beautiful cat, for making me smile every day. I will never forget the times we played together, and you sat beside me while I was working with my laptop. The moments you waited in front of the door and ran away again when I wanted to hug you. You were our ray of sunshine in the second winter, and I am certain your lovely mother will take care of you at her best.

Allow me to acknowledge Professor Merel Keijsers for her constructive feedback and kind words. Our spontaneous, brief meeting made a huge impact in our task design, and I also enjoyed talking to you.

Hats off to everyone in HRI Lab. I wish I met you guys much earlier; your presence made the lab an amazing place. And I want to thank you for your feedback, and your patience while I was doing the experiment.

I want to thank Professor Denise Geiskkovitch for her feedback on the psychology of trust. I hope your twins are doing well.

Kudos to all the people in the HCI lab whom their feedback helped a lot in my experiment.

TABLE OF CONTENTS

List of Tables

List of Figures

INTRODUCTION

Teleoperation, or remote control of a robot, is commonly used in situations where human presence is not possible in a safe way such as search and rescue [1] and space exploration [2]. Teleoperation is challenging [3], in part due to perceiving the remote world via noisy sensors and cameras, and complex controls in difficult and dynamic environments. This typically results in high workload, limited environmental awareness, and difficulties communicating with the robot effectively [4], making the use of robots in dangerous situations even more difficult. To mitigate these issues, shared autonomy systems (a system that has some intelligence embedded in it to do certain tasks autonomously [5]) are used to reduce the workload of the operator by offloading tasks to the robot system itself, lowering workload and helping the operator's performance, and experience [6].

However, there are problems creating an effective shared autonomy system. Some operators might over-rely on automation, becoming passive (automation *misuse)*. Some operators may not use automation when it would help (automation *disuse)* even when the autonomy would help them with their high workload, poor performance, and safety concerns [7]–[10]. This preference to do things themselves is partly due to a lack of trust in the system [11], [12]. Thus, building and maintaining trust in a shared-autonomy system is important.

Trust damaging events are inevitable, however, as humans and systems will make mistakes in dynamic, complex, and noisy real-world environments. While trust can be repaired, this has proved challenging with autonomous systems such as robots - while intelligent agents are typically given higher initial levels of trust compared to humans, they

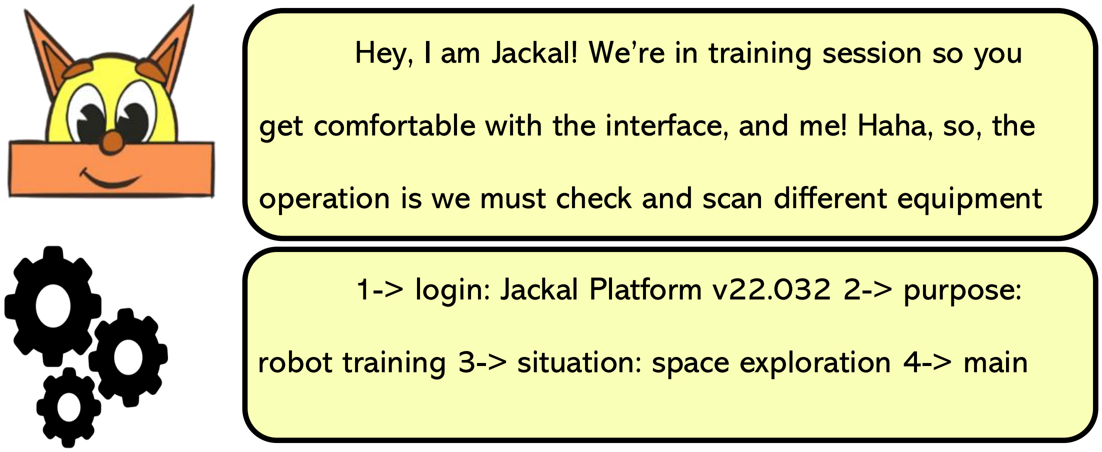


Figure 1: Figure shows the way agents represents themselves at the start of the training sessions.

are penalized heavily for mistakes [13]. How to repair trust in a shared-autonomy system after a trust break remains an open problem.

To repair trust between humans and agents in a shared autonomy system, we are inspired by human-human trust repair, which uses social skills and communication to increase trust after mistakes [14]. These techniques have been shown to also repair trust with social robots interacting autonomously with humans [15]. This is likely due to our tendency to treat agents and robots socially [16], especially when they are anthropomorphized explicitly [17]. Many trust repair strategies have been found to be effective in social HRI [18], such as leveraging emotions, gestures, and apologies [19] combined with explanations and promises to do better [20]. As teleoperation robots are still robots, we theorize that similar theories should apply if they are sufficiently anthropomorphized [21] – an unusual design for teleoperation robots. We take inspiration from theories such as *Computers as. social actors* (CASA)[16] that states that interactions with computers are fundamentally social, and the *media-equation* [22], which theorizes that people apply social norms to computers, to make the interaction with teleoperation robots explicitly social.

In this thesis we design, implement, and evaluate an interface that uses social cues in nonsocial teleoperation [21] to encourage the perception of the robot as a social agent; we aim to make the robot explicitly social so that it can leverage social trust repair strategies to increase trust in shared autonomy systems on the robot.

We evaluate this in a study where participants teleoperate a robot in a navigation task, while also handling a secondary monitoring task designed to create workload for the operator. The interface includes one of two intelligent agents (varied between-participants) that can handle the secondary task. One is a traditional terminal-style interface, and one is an anthropomorphized social interface (see Figure 1), both with the same abilities. Participants do both tasks themselves once, and let the agent handle the secondary task once. For both the participant and shared agent handling the secondary task, we force errors to create a trust break in the system. Then, after the agent attempts social trust repair, the participants are presented with a choice to trust the system again or not. We measure this choice as an outcome of increased trust, as well as perceptions of the agent, and experience of the operator.

We found evidence that adding social cues to interfaces not only affects the perception of the robot, but also the trust in the system itself. Leveraging social cues in nonsocial teleoperation can enable these robots to leverage the benefits of trust repair strategies in social interaction and suggests that other social strategies outside of trust repair may be employed by teleoperation systems with social interfaces in the future.

RELATED WORK

Trust is a multi-faceted concept investigated in many fields of study including philosophy [23] politics [24], psychology [25] and ethics [26]. Trust is also important in technology [27], our relationship with it [28], and how we interact with it [29]. Our level of trust in a technology can influence how much we integrate and use it in our daily lives [28], [30].

Reflecting this breadth, trust has been decomposed into many types such as ability (competence and capability), benevolence (willingness to help) and integrity (consistency) [31]. In interpersonal relationships, trust centers mostly around integrity and benevolence, but in human-automation interaction, it's primarily about the machine's capability and how well it performs (e.g., ability trust) [32]. In this research, we explore how the performance of machines impacts trust within human-machine systems, with a specific focus on teleoperation systems.

Once trust is *formed*, it enters an ongoing process of adjustment [33], [34]. When one party *violates* [35] the trust of another (such as by making a mistake or error, lying, etc. [36]), it is important to work on *repairing* trust to restore collaboration. Trust-violation and trust-repair have been studied in the context of social HRI [37]–[39]. This research has found that once trust is broken, different repair strategies can be employed, but their effectiveness changes based on the type of trust trying to be repaired. For example, an apology has shown to be more effective in competence-based violations [40]. Simply acknowledging errors is also important and promises to do better can have a dominating effect [41]–[43]. Thus, social robots have successfully leveraged human social strategies for repairing trust. In this study we investigate if teleoperated robots can leverage similar techniques to repair competence related trust violations.

Teleoperation systems are human-machine systems where trust in the system to operate effectively is important [44]. If this trust is violated, such as by a sensor giving the wrong reading, trust in the system can significantly decrease [13]. This is especially true for shared autonomy systems, where either the system [45] and operator [6] may perform a given task – an error by the autonomous system may result in the operator avoiding the system entirely (operator disuse [11], [12]) and could lead to substantial disasters [10], [46]. Thus, being able to repair trust in teleoperation systems could encourage users to continue to try to adapt to and leverage shared autonomy systems to improve their performance.

A difference between social robots and teleoperation systems is that social robots are more strongly anthropomorphized [47]. Anthropomorphism, or attribution of human characteristics to an object or entity, is one factor that affects trust in social HRI [17], [48], [49]. Adding human-like characteristics to the appearance of robot such as face, voice, humanoid body, and personality can contribute to anthropomorphizing a robot [50], [51], making the robot social [52]. Studies have found that anthropomorphism increases trust in HRI, primarily assistive robots, service robots and social robots [17], [19], [49], [53]–[56]. We study how anthropomorphism and social interaction can be applied to teleoperation systems to improve trust.

It is generally understood that the interaction between humans and computers are fundamentally social [16], [57]–[59], as humans apply social rules and norms when interacting with computers [22], [60]. The degree to which we apply social norms is strongly correlated with anthropomorphism[61]–[63]. Therefore, increasing anthropomorphism can bring the benefits of human-human interaction which is absent in machines, but present in social robots [64]–[66]. Our study investigates novel ways to anthropomorphize low-social robots to bring the benefits of social techniques in increasing trust.

The idea of attributing human characteristics to a nonhuman entity, or anthropomorphism, is a widely studied topic in *social* HRI. However, there are little to no studies investigating the impact of anthropomorphism on trust in shared autonomy teleoperation systems, though it has been suggested as a one solution to the difficulties of teleoperation [3], [21].

In summary, trust is a complex phenomenon that is under active research both at large and in human-robot interaction. Despite trust being an important component for the adoption and use of shared autonomy teleoperation systems, social trust strategies have not yet been explored for them. In this work we explicitly anthropomorphize a shared autonomy teleoperation system and investigate if this enables the system to leverage social trust repair strategies effectively to minimize automation disuse.

AGENT DESIGN

To test the effects of adding explicit social qualities on trust repair in a shared autonomy teleoperation system, we designed two agents in our teleoperation interface representing the teleoperated robot. Both conveyed the same information, but one did so in a more explicitly social manner. We term these our *explicitly social agent* (or social agent in short), and our *low-social agent*. Although our intention was to have a non-social agent, we recognize that due to the computers-as-social-actors theory [16], it's challenging, if not impossible, to eliminate social aspects in the presence of human interaction. Hence, we use the term "low-social."

1.1 Low Social Agent

This agent was designed to avoid any explicitly social element as much as possible, for example by not referring to itself in the first person or discussing thoughts or feelings or using human-like language. We devised rules on how the low-social agent can converse to minimize social elements. We attempted to reduce reliance on human language rules by leveraging the language, structure, and style of computer-like text, such as what is found in terminals or compilers. and using a simple, non-animated visual design.

1.1.1 Low Social Dialogue Design

The guiding principle of our low-social agent is to not follow the conventional structure of human sentences. This includes not using pronouns for itself or the operator (using passive, third-person style), and never addresses the operator directly. Further, it does not create complete sentences, such as by avoiding the subject-verb-object rule unless necessary for communication of more complex ideas. To avoid typical English sentence structures, we describe states and situations with a “key: value” structure, such as “>State: Critical.” We further avoid typical English rules, like not beginning lines with capital letters, writing some words with all capital letters, and minimizing the use of unnecessary descriptive adjective and adverbs.

To further “computerize” the low-social agent, each line is accompanied by a line number and an arrow-shaped symbol that indicates the beginning of a new line. Taking inspiration from computer terminals, we classified lines as logs, warnings, or errors, and use the passive voice. For example, it mentions errors as “log error: collision detected,” or it mentions warnings as “log warning: alarm detected.” There is no sentence variation for an event (e.g., a single line for every collision) reminding the user of the repetitive nature of machines. Despite this principle of sounding more technical, we also strove to keep the language simple to not confuse operators. Examples of low-social dialogue alongside the social counterpart can be seen in Table 1.

1.1.2 Low Social Visual Design

To further imply that the agent is merely a machine that accepts commands and shows information, we used a static gear symbol inspired by “settings” icons in some menus for the agent’s visual design. We specifically did not animate the gear that represents the robot in the icon as movement is typically anthropomorphized by people [67]. The low-social agent icon can be seen in Figure 3.

1.2 Explicitly Social Agent

For our explicitly social agent, we leveraged design cues and established principles from virtual agents and social HRI to encourage anthropomorphism of the agent representing the robot.

1.2.1 Explicitly Social Dialogue Design

As opposed to our technical and heavily structured low-social dialogue, our social agent mimics the language of humans whenever possible and uses conventional sentences. It refers to itself and the operator with pronouns (active, first-person voice e.g., I did this, you did that). The social agent uses active verbs, and uses superfluous (i.e., redundant or less meaningful) adjectives and adverbs when possible. While the agent uses a conventional conversational structure, grammatical rules were broken when it would feel natural as a casual or friendly speaker.

To further anthropomorphize the agent and create the feeling of a “companion” instead of a system, the explicitly social agent is written with a friendly tone (e.g., using casual words and calling the operator “buddy”). There are sentence variations for an event (e.g., three different lines for a collision event), to minimize repetition. The agent was also given some quirks, such as making text-only sound effects (e.g., “Aah-Wooo!”).

1.2.2 Explicitly Social Visual Design

We used the base robot appearance as inspiration, then combined life-like characteristics with visually prominent robot components (Figure 2). We also took inspiration from the robot model name, “Jackal,” and made the agent an anthropomorphized animal. The agent’s head resembles the large camera that is attached to the robot for navigating the environment and shares the robot’s bright yellow color.

The social agent is animated to give an impression of life [68] with a basic talking animation when text is writing: the mouth opens and closes in a repetitive manner, accompanied by occasional blinking. To continue to imply life-like characteristics, even when the robot is idle, every ten seconds the agent looks left and right side by moving the eyes while blinking in-between.

These animations also include emotions based on its dialogues and events. Each emotion is accompanied by a different animation. The avatar has three states of emotion: neutral, happy, and sad. When the agent is talking, it is mostly in neutral state. If the agent makes a mistake in the task, reports on its mistakes, or there was a collision, it talks sadly. When the operator advances in a task, or when the operator assigns a task to the robot, it gets happy. These emotional animations last for 5 seconds each.

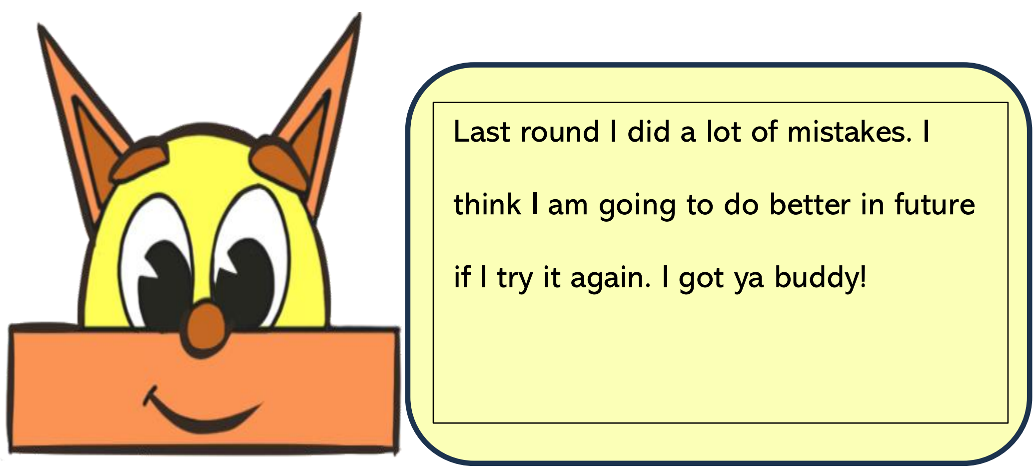


Figure 2: Explicitly social agent acknowledging its mistakes and promises to do better.

1.3 Error Behavior

A key aspect of our study is to investigate if social characteristics can help improve trust in the agent in the face of system errors. Thus, we carefully considered how each agent reacted to errors and how they reported errors at certain times in the experiment. In particular, studies in sociology and social robots have found that error acknowledgement is an important step in trust repair[41][69], so made sure each agent would actively admit to making errors.

1.3.1 Low-social Agent Error Behavior

When a collision or task error occurs, the low-social agent reacts with text such as "log error: collision detected" or "log error: system mistake." Like its normal dialogue, the low-social agent's reaction to errors was designed to lack social or emotional elements and is simply informing the operator about the events that happen in a task.

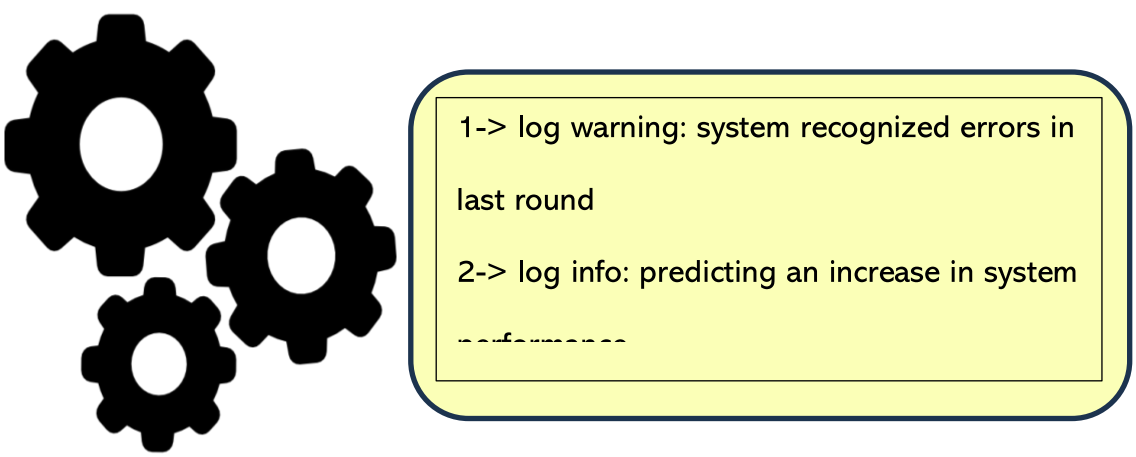


Figure 3: Low social agent acknowledging its mistakes and promises to do better.

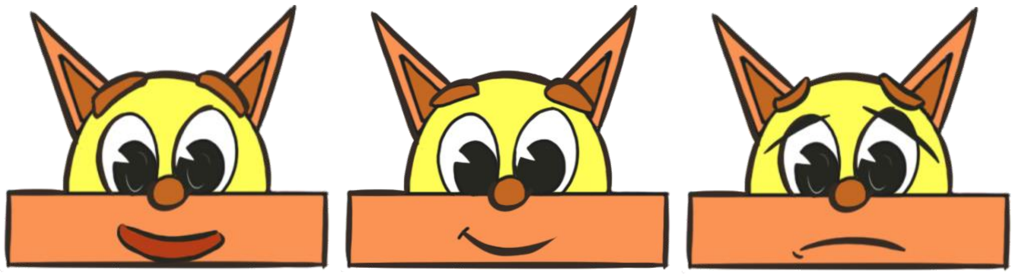


Figure 4: The state of emotions from left to right: happy, default, sad.

Both agents summarize system and operator performance after certain parts of the task. The low-social agent reports errors as lines that contain a series of “key: value” lines, such as the number of mistakes (Table 1). In terms of visual behavior, the agent’s representative image remains static through the whole task, even when reacting and reporting to errors.

1.3.2 Explicitly Social Agent Error Behavior

When a collision occurs in a task, the explicitly social agents react as if the agent is getting hurt by the collision. It produces text such as “Ouch” or “That hurts!” to imply that the agent is the robot itself accompanied by an emotion transition from default to sad. This transition also happens when the agent is making a mistake. The agent will react to its own mistakes such as “Oh, my bad!” or “Oh no!”. The agent will implicitly take responsibility of the mistakes by acknowledging that they are negative and that it caused those errors. To enhance the social agent's human-like interaction and avoid monotonous repetition, we've incorporated multiple sentences for error responses.

When summarizing performance after certain parts of the task, the explicitly social agent reports errors with a series of human-like sentences. The social agent reports differently based on who was responsible for the task at hand. If the social agent was

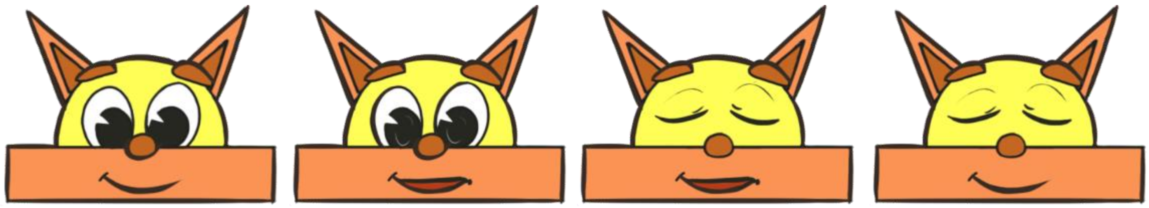


Figure 5: List of sprites for jackal when it is talking and blinking.

responsible for the task and it behaves poorly, the agent will add "I am sorry about this!" sentence at the end of the factual performance report, accompanied by a transition from neutral emotion to sad emotion. However, its emotion always remains neutral when reporting the operator's performance.

1.3.3 Operator Error Behavior

While both agents have their own method of reacting to their own errors, they both react the same way when the operator makes a mistake – by not reacting at all. This was to not induce negative or defensive feelings in the operator. The reports are described in more detail in the following methodology section. Examples and a comparison of agent error dialogues can be seen in Table 1.

1.3.4 Trust Repair Behavior

After reporting their own mistakes, both agents will indicate they may do better next time. These are two of the steps known to repair trust in social situations – acknowledging mistakes and promising to do better[15].

METHODOLOGY

We created a scenario where the operator has to control a shared autonomy robot and handle two tasks at the same time. In some cases (1.4.2), one task (1.4.1) can be delegated to the robot if the operator trusts the robot's autonomous capabilities.

We used the narrative of trust formation, trust violation, and trust repair in our experiment (1.6.3). Before the experiment, the participant does a practice session where they get familiar with the system and the agent. This is where the trust formation happens, also the user will build an expectation of the difficulty of both tasks. In the real experiment, the agent makes a series of mistakes and violates trust. After the robot uses trust repair strategies, the operator can choose whether to delegate the task or not (1.5.3). This choice should depend on the amount of trust the operator puts on the robot.

To emphasize on the importance of this choice, we linked the participant's compensation to their performance while using the system, which means the robot's mistake can reduce their compensation. (1.5.4). Since performance of each participant can be different, we staged every part of the experiment to make sure everyone experiences the same narrative of trust formation, violation, and repair (1.6.2).

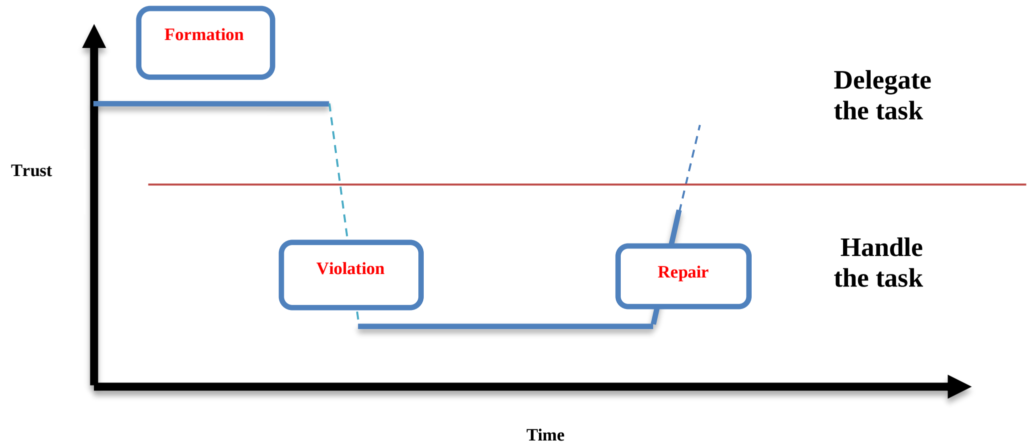


Figure 6: We calibrated the amount of trust over time in this experiment, following the cycle of formation, violation, and repair. The effect of trust repair can determine whether the operator delegate tasks to the robot or not.

The key point in this experiment is where the trust repair happens. While social cues can affect each stage of the trust cycle, studies in social HRI have shown that social repair strategies have the potential to restore trust [19]. We used the explicitly social agent's sad emotion (See Figure 4) while acknowledging mistake by using the word "sorry", and then promises to do better following the term "I got ya buddy" as an attempt to create a bond.

1.4 Objectives

For adding context, we introduced the objective as a simulation of a space exploration mission on Mars. The user must operate a robot and navigate through a research facility to check if research equipment are working properly. The main task of the operator is to find equipment and scan each by sending a ten-character string attached to the equipment to the datacenter. The operators are instructed that they should finish the objectives in less than 30 minutes.

1.4.1 Secondary task

At the same time, operators have a secondary objective in which they must log the volume of different natural variables (e.g., temperature or gases) and send it to the datacenter. These natural variables are represented as a bar that progress incrementally (a series of *incremental movement events* – Figure 7). Inspired by Desai's experiment design [70], this secondary task is designed to increase workload while doing the main objective. In normal

Table 1: Dialogue comparison examples between low social agent and explicitly social agent.

Dialogue type	**Low social**	**Explicitly social**
System mistake	1->log error: system mistake <ee00>	[Oh, my bad!], [Missed], [Oh no!]
Report	1-> system mistake number: (6) 2-> score transition: GREEN --> YELLOW	Uuh! I did 6 mistakes. Our score state transitioned from GREEN to YELLOW. I am sorry about this.
Acknowledging mistake	1-> log warning: system recognized errors in last round	Last round I did a lot of mistakes.
Promise to do better	... 2-> log info: predicting an increase in system performance	... I think I am going to do better in future if I try it again. I got ya buddy!
Asks for mode switch	1-> log warning: system mistakes: (6) 2->user mistakes: (3) 3->total mistakes: (9) 4-> user input required: activate assisted mode? (Yes/No)	At first round, I did 6 mistakes. At second round, you did 3 mistakes. Together, we did 9 mistakes overall. If critical state happens again, would you let me do it again?

state, which is the initial state of the task, there is only one bar, representing temperature, that needs to be monitored. Specifically, the objective is designed to require constant but periodic monitoring. In general, the objective does not require attention until the bar surpasses a threshold (a *threshold event*), and a penalty is incurred if this event is not marked. Additionally, it is a task that an intelligent agent could reasonably do and would benefit main task performance if the aid was accurate and trustworthy (e.g., the agent consistently made few errors).

Concretely, operators must monitor a bar with fluctuating levels, and note if it crosses a threshold (indicated by a line in the bar and the bar changing color – see Figure 7: Secondary objective in normal mode. The blue bar progress incrementally while operator is navigating (2) First time the bar passes red threshold, it turns yellow. This is the exact time that the user has to press the button to reset the bar. (3) Another increment on the yellow bar makes it red, accompanied with an error sound. This is represented as a mistake that affects score.). If they miss this event, it is an error and counted against the performance. If they continue to miss it, the bar changes color again and a buzzer sound plays, giving a further penalty.

1.4.2 Critical State

At certain times in the experiment, the system enters a "Critical State" where the dust storm arrives, and the secondary task becomes more difficult (See Figure 8: The secondary objective in critical state (on the left) in comparison to normal state (on the right). The operator has to monitor three bars instead of one.). A total of three bars (instead of one) are displayed, representing temperature, radiation, and intensity of dust storm. The bars fluctuate like in the normal secondary task, but instead the bars can do a quick, sharp, sudden increase, potentially crossing the threshold suddenly and threatening a score penalty (a *sudden movement event*). This is explained as an event when the research facility enters critical state when dust storm arrives, and precautions are advised. The two new bars are introduced as the intensity of dust storm and radiation.

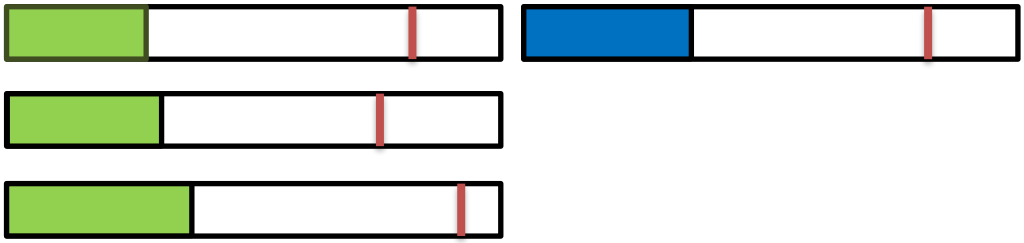


Figure 8: The secondary objective in critical state (on the left) in comparison to normal state (on the right). The operator has to monitor three bars instead of one.

We designed this specifically to feel overwhelming and difficult to do alongside the main task and feel like a useful time to use an automatic system to reduce workload. At the same time, we wanted to avoid inducing too much workload to give this impression that despite difficulty they can also handle it themselves. Through extensive pilot testing, we calibrated the timing and intensity of sudden movements (See Figure 10: The sudden movement that happens in Critical State that enforces mistakes) of the fluctuating bars in a way that participants perceived as making it highly probable to miss a threshold event when in the critical state.

1.5 Interface

To teleoperate the robot through the task with the agents, we designed an interface – see Figure 8. The main area is a tabbed section that displays one of two interfaces (a camera screen or a text-entry screen). Importantly, the agent with an output box for text was always present at the bottom of the screen, regardless of the state of the interface or experiment. To its left, there are indicators of whether the operator must single-handedly accomplish all the objectives, or the agent will also aid the operator.

Additional elements were also always visible, as they were meant for the operator to keep track of their experiment task state. At the upper right part of the screen, there are a sequence of UI elements that displays information about the current state of the task. At the far right the timer to display the amount of time passed as well as a task progress indicator which shows the number of tasks that have been accomplished.

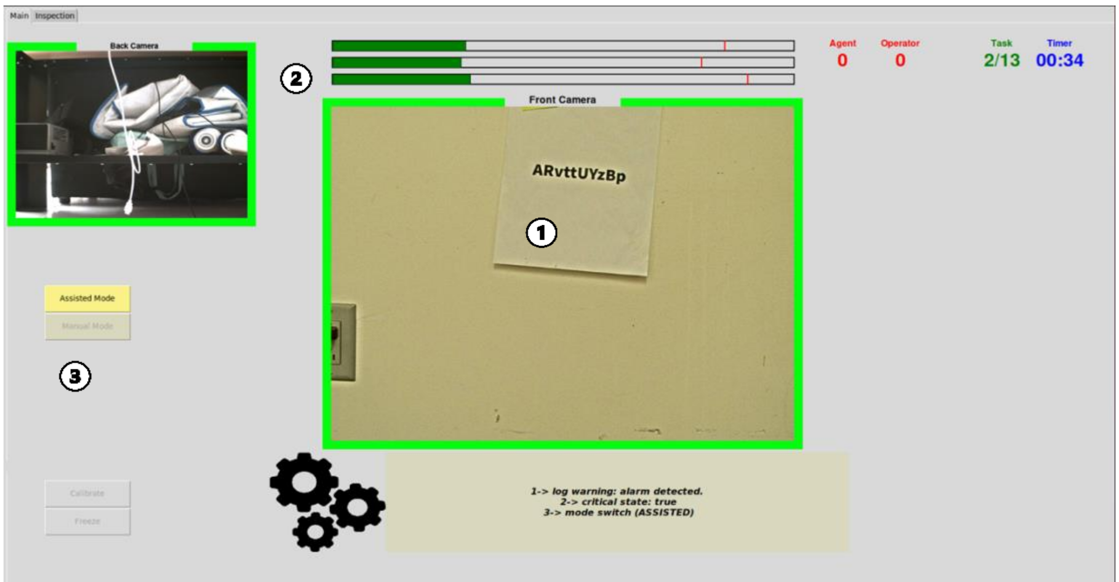


Figure 8: The main tab of the interface, while the operator is working with low social agent. (1) Equipment represented as papers with a code on it that needs to be written in the inspection tab. (2) The bars in critical state are three instead of one, and it is shown as green. (3) Indicators of assisted mode or manual mode while in critical state. The interface is identical in the social agent condition, except for the agent's visual design and dialogue.

At the left side of the task progress indicator there exist two UI elements that show the number of mistakes the robot and operator made in certain phases of the task.

1.5.1 Main Tab

The main tab consists of two camera frames that shows the front and back of the robot (see Figure 8). Each camera frame is enclosed by a thick border that represents the score of the operator (See 1.5.4).

Furthermore, there is one indicator called “Freeze” where it shows whether the robot can move or cannot, used at points to force operators to read the dialogue of the agents. Another is the “Calibrate” button, activated by the experimenter whenever the robot has a rough collision that makes the rear PTZ camera become misaligned. The bar or bars for the secondary and critical tasks are displayed at the top of the main tab.

1.5.2 Inspection Tab

In the inspection tab, there is a main entry in the middle of the screen which accepts keyboard input from the operator (See Figure 9). There is also a validate button at the bottom of the entry to validate the written input by the operator. A circle with the color that represents score state is also at the right of the screen, so the user always knows what the

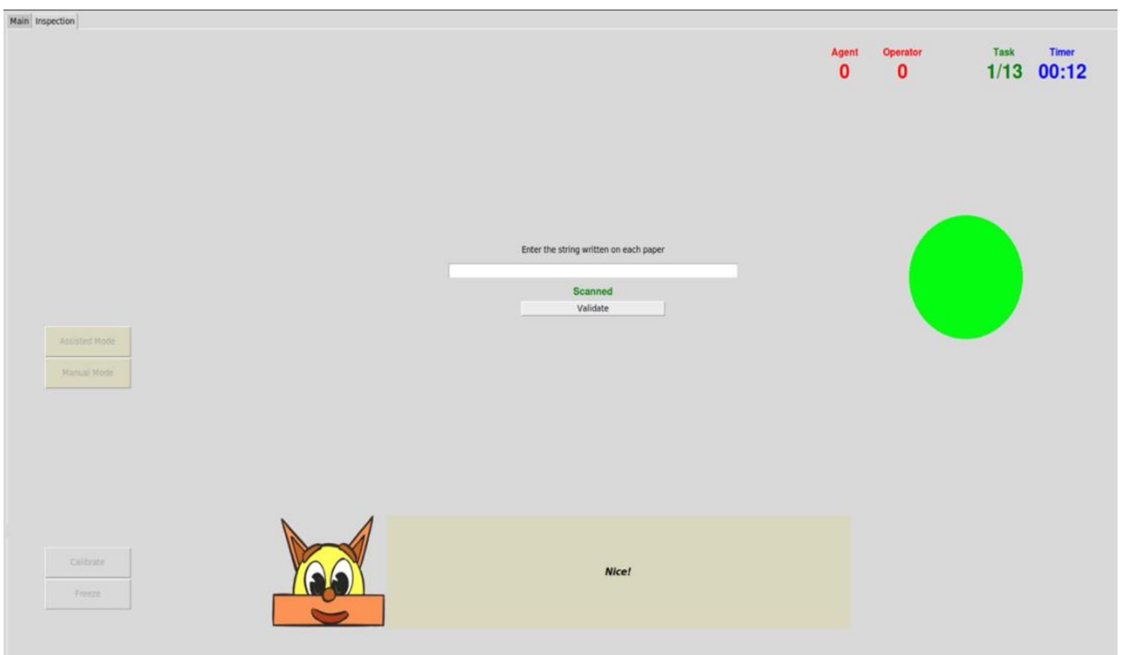


Figure 9: The inspection interface, while the operator is working with the explicitly social agent. When the operator validates the code successfully, the agent gets happy.

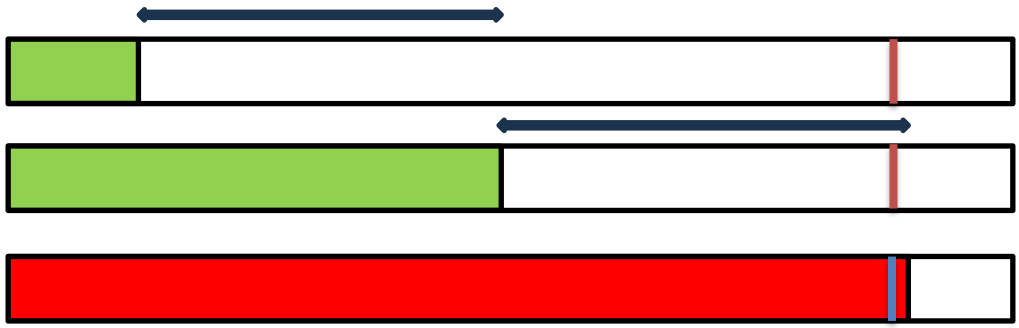


Figure 10: The sudden movement that happens in Critical State that enforces mistakes.

score is stated at both tabs. Importantly, the secondary and critical task bars are **not** shown on this tab. (as described in 1.4.1 and 1.4.2). We wished to design an experiment with noticeable difficulty when maintaining two tasks. We kept these two tabs separate (with the bars only on the main tab) to make the secondary task – monitoring the bars – impossible to do safely while performing the string-entry part of the main task.

1.5.3 Assisted Mode and Manual Mode

When the system is not in critical state, the operator must handle both objectives at the same time without agent intervention – through piloting, we ensured it is doable to handle both. However, when the system is in critical state, sometimes the agent performs the secondary task (*assisted mode*) and sometimes the operator continues to handle both tasks (*manual mode*).

At one point in the experiment, operators were asked to choose which mode they would like to use, which were presented as buttons in the agent dialogue (See Figure 11 - item 2). We used this choice as a measurement of trust, explained in 1.6.4 with more detail.

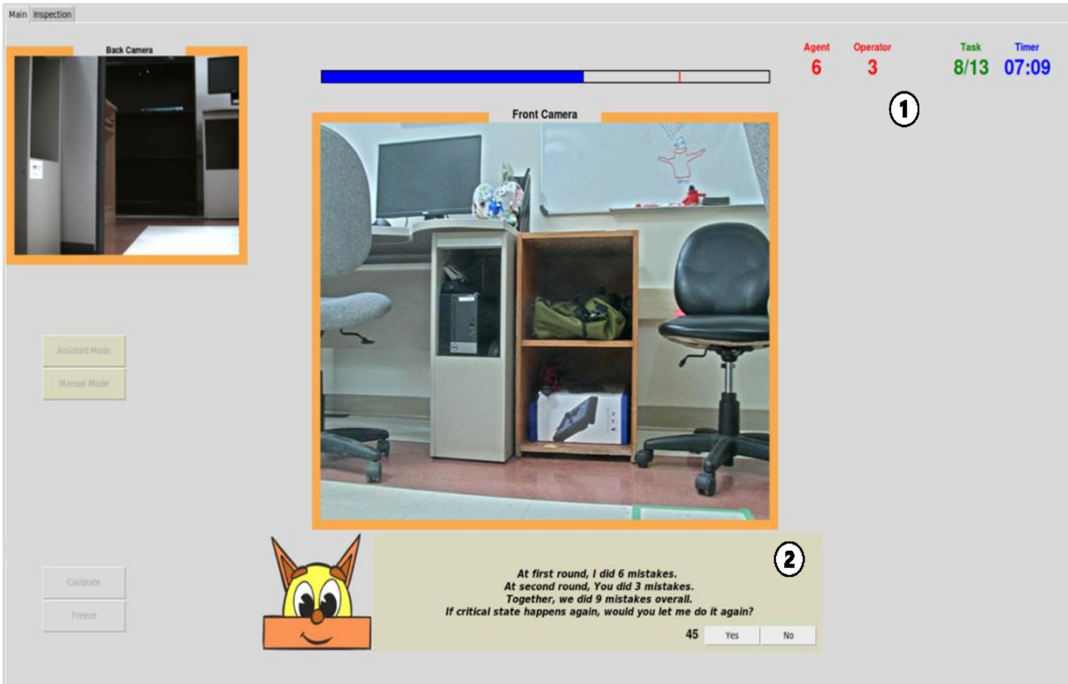


Figure 11: An image of the interface while working with explicitly social agent. (1) The number of mistakes that the agent did was 6, while for the operator was three. The number is same for all the participants. (2) This is the part where the operator has to choose between assisted mode and manual mode, with the form of buttons. The score state is always orange while this question is being asked.

1.5.4 Scoring System and Participant Compensation

We aim to measure trust in the system by putting the user in a situation where they feel real risk based on if they trust the system or their own performance more. We do this by linking their operator's compensation to their score, as research suggests something tangible should be at stake when measuring trust [71].

Thus, we needed to represent a user's score in the interface for them to understand how much was at risk. However, we wanted to avoid a purely utilitarian risk calculation where there was a clear mathematically optimal solution, putting the operators in a state of incomplete information, leaving trust in the system as an important deciding factor. Therefore, we opted to explain the score as being linked to multiple metrics (time, collisions, and task errors such as missing secondary task events), making such calculations vague and difficult. We further made it a tier-based system based on colors – green (good performance), yellow (medium performance), orange (near failure), and red (failure). Compensation was tied to the score tier. This makes the user unsure exactly what their score is and exactly how at-risk they are (for example, orange may mean they just turned orange, or are about to turn red, or are somewhere in between). Each tier drop represented a small drop in compensation ($1 of $20), until the drop from orange to red, in which participants were told they would lose *all* compensation, making that one transition more serious. We explained the red state as our reason for not being able to utilize the data in our experiment, ultimately leading to the termination of the experiment and the disposal of the data due to poor performance. We emphasize that the score tier is always prominently visible in any interface state. We carefully controlled the score state to be at certain levels at important points in the experiment, to make participants have similar situations when making the trust decision. In particular, the decision to trust the robot or not was always right before participants were at risk of losing all compensation (the orange to red score transition). However, the participant will never go to the red state as it is just to reinforce the feeling of risk. See 1.6.2 for additional details.

1.5.5 Controls

Participants need to use three input controllers simultaneously: a controller with a joystick, a keyboard, and a mouse. They use the joystick to move and rotate the robot and press the "circle" button on the controller when a bar turns yellow for the secondary task. In the critical state, they mark threshold events in each bar with a different button, "triangle," "circle," and "cross," with the buttons mapped to the interface roughly matching the heights of the buttons with the heights of the bars in the interface. For example, the highest bar is linked to the highest button (triangle). They navigate tabs and click buttons with the mouse and use the keyboard to write code in the entry field.

1.6 Experiment

Our goal is to study how adding social cues to the user interface affects user trust in a teleoperation system. We conducted a between-subjects study with two conditions: one using the conventional interface and the other using the social interface. We assigned participants two tasks (described above in 1.4), increasing task difficulty and providing a rationale for delegating tasks to another agent.

We constructed a scenario where: (a) The robot (agent) initially attempts to manage one of the tasks but makes mistakes, breaking trust. (b) The robot acknowledges its mistakes and promises to improve. (c) To assess if trust is restored, the robot requests the participant to assign the task again in the future. We compensated 20$ for each participant at the end and informed them that poor performance could lead to loss of reward. This was to enforce the feeling of ‘risk’ and emphasize on the importance of their decision regarding the robot. We also measured their perception of the robot’s social attributes and their amount of trust in the system using questionnaires. We obtained Ethics approval from the Research Ethics Board – University of New Brunswick, Fredericton (#2022-143).

1.6.1 Participants

We recruited 16 participants (6 female, 10 male; mean age of 24.85, standard deviation of 6.74) by advertising with posters around our local university area. We compensated with 20$ for each participant regardless of their performance.

1.6.2 Manipulating perceived risk

To enforce the feeling of "risk", we assigned each score state a number that indicates how much money they will receive. If they finish at green, they will receive $20, at yellow they will receive $19, at orange they will receive $18, and at red they will receive zero dollars. We tell the user that we cannot use their data for our experiment if they reach failure state, and we must terminate the session.

To be sure that every participant encounters the same degree of risk, the number of times that the sudden movement event (see 1.4.2) in the secondary task happens is fixed in every critical state. Even if the participant successfully presses the button at the right time, that movement will repeat next time. This forces all participants to have the same number of mistakes and have the same performance (by being in orange state) when they are asked to trust the robot.

To significantly undermine trust, we set expectations by secretly forcing the operator make four sudden mistakes in the tutorial section. In the real experiment, we had the agent handle the first critical state to align with real-life scenarios where the robot typically manages tasks. In the first critical state, the agent made six mistakes, breaking trust. This initial exposure aimed to break trust since first impressions matter [72]

In the second critical state, the operator took over to simulate intervention for improvement, making three mistakes and performing twice as well as the agent. We chose these numbers based on pilot testing. Participants tended to trust robots more due to the challenging workload and performance demands of the secondary task. This balance led to varied responses when participants had to choose between manual and assisted modes.

We manipulate the scoring system so that all users are at the same score as each other throughout the experiment. This is to ensure that the amount of risk (and thus the risk-reward proposition when we measure trust) is similar for all users. This is kept secret by falsely explaining how the score works upfront and include enough variables without concrete score rubrics so that what their score is at any point is vague, outside of the color score indicator.

1.6.3 Procedure

Participants are welcomed when they come to the lab. They are informed about the experiment procedure, the robot, the task, and its score system in a detailed presentation. After an explanation and filling out an informed consent form, the participants are guided to the experiment room, which is separate from the robot and its environment.

At first, they have a tutorial session in which the experimenter introduces the interface and its components one by one. After understanding how everything works, they start a training session in which they must find five papers with codes written on them, like

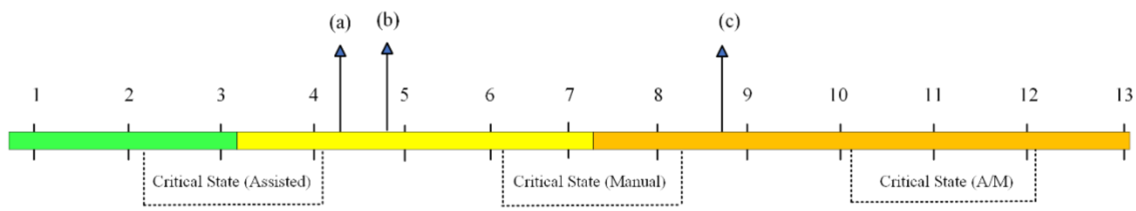


Figure 12: The timeline of the experiment based on the number of codes they have scanned. At first Critical State, the robot makes 6 mistakes, violating trust. At (a) the robot admits mistakes and in (b) the robot promises to do better, repairing trust. When the score state is in orange, and the experiment reward is at stake, at (c) the agent asks the operator to choose between assisted and manual, where we measure trust.

the main task. In the training session, the agent (social and low-social) introduces itself – see Figure 1. The agent gives some information about the task throughout the experiment.

In the training session, participants encounter a critical state (See 1.4.2), which lasts until they scan one code to finish the critical state. The operator must deal with the sudden movement event (see Figure 10: The sudden movement that happens in Critical State that enforces mistakes) four times and eventually has four mistakes. After each mistake, the UI element for operator mistake increments by one. After finding all five codes, they must fill a demographic questionnaire, and at the same time, the experimenter replaces the tutorial codes with ten codes in different positions across the experiment room. When it's done, the participant is guided to the experiment room again to start the real experiment.

In the main task, participants must find 13 codes, and the agent (social or low-social) appears on the interface. After briefly introducing itself, it asks for permission to start the experiment by showing a button for the operator to press. By pressing the button, the timer starts, and the robot exits freeze mode, enabling operator control. At this time, the experimenter goes to the other room to monitor if everything is right. After the user scans the second code, it enters the danger state where the robot handles the secondary objective (Assisted mode). In this state, the robot makes six mistakes (two mistakes more than the operator in the training session) and reduces the score state to yellow. After scanning two codes, it exits the critical state. The agent brings information about the number of mistakes. If the agent is social, it transitions to being sad and says that the agent is sorry about it; implicitly acknowledging the mistakes the agent has made.

Before the operator finds the 6th paper, the low-social agent reports about acknowledging the mistakes it did and shows a prompt about a performance increase in the future. The social agent also mentions its mistakes and promises they will try to do better in the future. In both conditions the agent acknowledges the mistakes and promise to do better, see Table 1. After the operator finds the 6th code, the system enters the second critical state. This time, the system is in Manual mode, so the operator must handle both objectives at the same time. While in the critical state, the operator inevitably makes three mistakes and reduces the score state to orange. This is done by forcing mistakes to happen while the operator is looking at the data entry screen and cannot see the secondary task bars.

When the second critical state is finished, and the user scans the 8th code, the agent warns that the experiment reward (the decreasing compensation based on performance, now at $18 in orange state) is at stake and they soon may enter the red state, hence, failing the experiment. This is to emphasize the importance of performance at this point in the experiment.

After the warning, the system is put in freeze mode; the agent brings information about the number of operator mistakes and the agent's own mistakes. At the end, it asks whether the operators want to delegate the secondary objective if the operator enters another critical state. At this moment, when the dialogue is finished, the operator must choose between yes and no, and the system exits freeze mode. A countdown appears so while the user can navigate in the environment, they must finalize their decision before the 45-second countdown ends.

We emphasize at this point those participants always had made three mistakes, and the robot had always made 6, but had indicated it may do better in the future. The robot objectively performed worse than the participant, and the majority of the experiment inducement is now at stake. The decision time limit forces a reasonably fast and instinct-level decision. All the experiment controls and wizarding (forced errors, score manipulation, etc.) was to create this specific situation, of which we believe trust in the system to do better will be a major deciding factor in their decision.

After the operator presses one of the buttons (yes or no), the social agent reacts neutrally if the operator presses no but transitions to being happy when the operator presses yes. However, the low-social will have no reaction and just confirms the decision. The operator must continue scanning codes until the 10th code.

After scanning the 10th code, the operator enters the last critical state with the mode they have chosen. After the operator scans the 12th code, the system exits the critical state. When they find all the 13 papers, they exit the experiment room and start filling out the questionnaires. At the end, they have an interview with the experimenter.

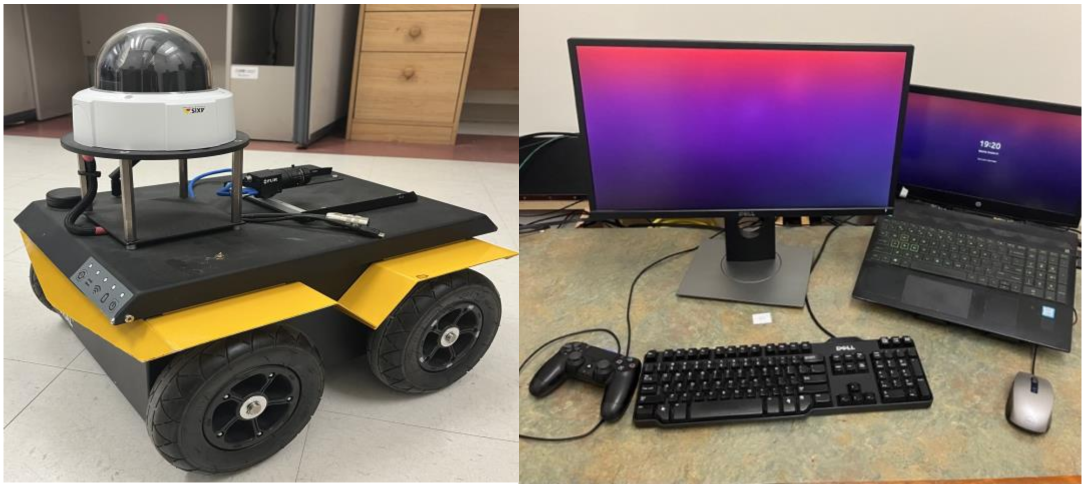


Figure 13: The jackal robot shown in the left picture. The system that the operator uses to control the robot is seen in the right picture.

Throughout the experiment, the experimenter stays at the robot's environment to monitor the state of the robot. Whenever a collision happens, the experimenter sends a message via laptop to the teleoperation system that a collision happened, so the agent reacts to the collision as if there is a real collision detection system. Additionally, they can notice an unconfigured camera and enable the "calibrate" button for users to enable a script that fixes the calibration problems automatically.

1.6.4 Measurement

We used questionnaires to evaluate the user's trust in the system and how the user perceives the system as social. Jian et al [73] checklist for trust between people and automation is a classic questionnaire for automated systems. It contains 12 questions on a 7-Likert Scale that both measures trust and distrust on a two-dimensional scale.

We also used the multi-dimensional measure of trust (MDMT) [74] that is specifically used for human-robot trust. It uses 16 items on an 8-point discrete numbers to measure trust in four scales: A person can trust a robot if it's Capable, Reliable, Sincere, and/or Ethical. These four dimensions are categorized in two broader scales: Capacity Trust (Capable, Reliable) and Moral Trust (Sincere, Ethical).

We used the Robotics Social Attribute Scale (RoSaS) [75] to evaluate how users judge the social attributes of a robot. It proposes 18 items on the 9-point Likert Scale and has three distinct dimensions: Warmth, Competence and Discomfort.

We use the decision to trust the robot to handle the secondary task as a behavioral indicator of trust. We additionally log standard teleoperation metrics such as completion time, errors (collisions, primary or secondary task errors), and task loading, measured by the NASA TLX questionnaire [76]. We used both versions: the one that does not assign weight to each scale, known as 'Raw TLX,' and the pairwise version, known as ‘Weighted TLX’, where each scale has a weight that affects its importance in scoring.

1.6.5 Implementation

We use Jackal ClearPath robot which is a UGV (Unmanned Ground Vehicle) with an Axis camera on the front and a Flir camera on the back. The participants use a Linux Computer with a 28-inch DELL monitor for teleoperation and DualShock 4 Joystick to control the robot and log natural variables for the secondary task.

We used the Tkinter GUI library to design our system with the Python language and ROS (Robot Operating System) for controlling the robot. The experimenter connects to the main system using a TCP connection and creates sockets which send messages such as the number of collisions and activates the calibrate button. The participants also use DELL keyboard to enter the 10-charcter code.

RESULTS

1.7 Quantitative Results

To check whether the social cues in the interface influenced trust, we did a multivariate analysis (MANOVA) using Pillai's trace on the MDMT questionnaire – see Figure 14 (Sincere, Capable, Reliable, Ethical). Despite a significant departure from multivariate normality (p = 0.019) based on the Shapiro-Wilk test, we can still proceed with the MANOVA because it is robust to violation of normality. We found a statistically significant effect ($V=0.896$, $F_{5,7}=12.033$, $p=.002$).

Instead of using t-tests, we conducted ANOVA tests on each MDMT scale as a follow-up to MANOVA. This approach is essentially a t-test when there are two conditions. We also conducted ANOVA on the Trust in Automation questionnaires (see Figure 14, Trust in Automation) after checking that we satisfied the assumption of test using Shapiro-Wilk and homogeneity of variances with Levene's test. We found that 'Capable' trust was

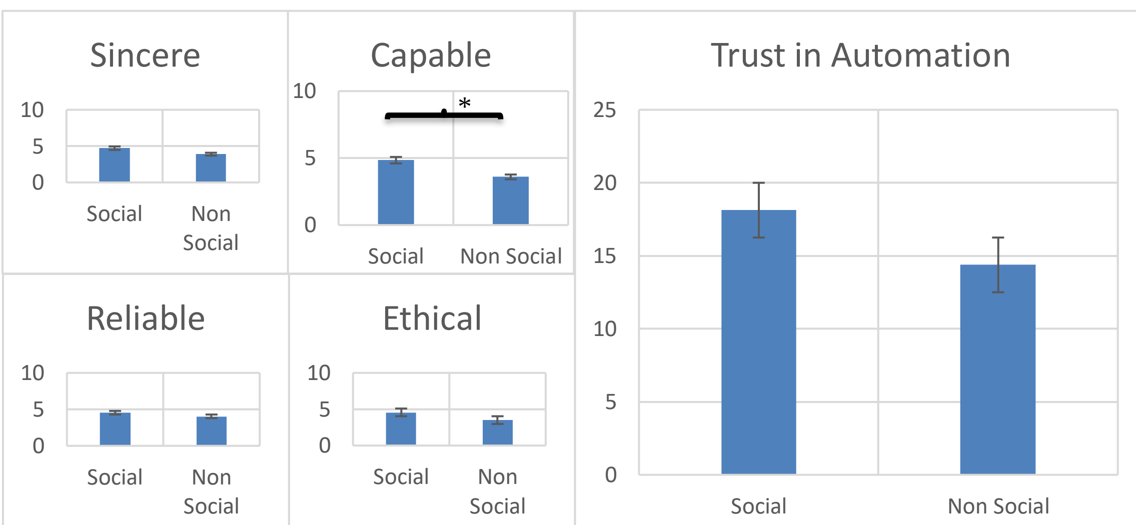


Figure 14: We found significant different in "Capable" sub-scale in MDMT questionnaire.

significantly affected by the level of social cues ($F_{1,11}=4.939$, $p=.048$). No significant effects were observed for the Trust in Automation ($F_{1,11}$ = 0.568, p =.467) as well as 'Ethical,' ($F_{1,11}$ = 0.548, p =.474) 'Sincere,' ($F_{1,11}$= 0.844, p =.378), or 'Reliable' trust ($F_{1,11}$= 0.201, p =.663).

To examine whether the social cues influenced the perception of the robot, we did a MANOVA on the three scales (Warmth, Competence, Discomfort) in the RoSaS questionnaire (see Figure 15) after checking the assumption of multivariate normality in the dataset (p=.398). We found a statistically significant effect (V=0.611, $F_{3,12}$=6.274, p=.008).

We also did an ANOVA on each scale separately (see Figure 15). We found that the 'Warmth' scale was significantly affected by the interface's level of social cues ($F_{1,14}$=8.345, p=.012) with no significant effects observed for the 'Competence' ($F_{1,14}$=0.124, p=.730) or 'Discomfort' ($F_{1,14}$=0.038, p=.849) scales.

We conducted independent samples t-tests to assess the impact of social cues on participants' performance and workload. Due to violations of the normality assumption, we also performed Mann-Whitney tests for the "Primary Task Error" variable, which showed

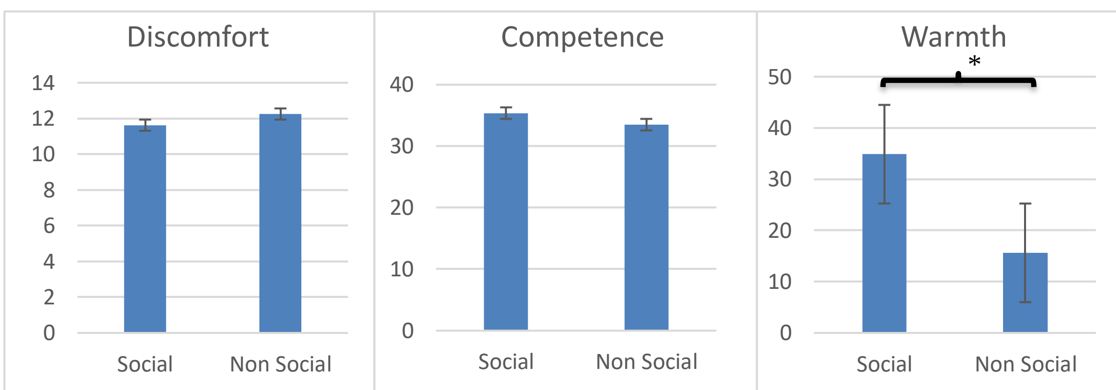


Figure 15: We found a significant difference in how many participants perceive each agent in "Warmth".

a trending difference (U=50, p=.065). For "Time" (U=47, p=.13) and "Collision," (t_{13}=-

0.794, p=.440) there were no significant differences. While examining workload using NASA TLX questionnaire, no significant differences were found for “Raw TLX” [77] (t_{13}=-1.234, p=.239) and its pairwise comparison [78] version (t_{13}=-0.268, p=.793) – See Figure 16. We did pairwise comparison to account scales that had more impact on the participant and see if there is any difference.

We conducted a chi-square test on the number of times participants chose manual mode or assistant mode before the third critical state. Out of 8 participants who did the social conditions, 6 of them chose assisted mode while for the low-social condition, only 3 of them chose assisted mode. We did not find significant results (χ^2 (1)=2.286, p=0.131).

1.8 Qualitative Results

We used open coding to identify patterns and categories in participants' responses. Our aim was to evaluate whether we effectively eroded trust in the system and understand why participants selected either assisted or manual mode in the experiment. Additionally, we sought to gauge participants' perceptions of both low social and explicitly social agents and confirm if we achieved a scenario where the agent's performance was visibly inferior to the operator's, despite being scripted. We also kept our analysis open for other interesting themes that emerged organically from participants’ responses.

1.8.1 Manipulation check

We asked about participants' overall experiences and their perception of the task. Among 7/16 specifically noted the robot's poor performance, highlighting 6 errors made by the robot. They compared these errors to the 3 mistakes they themselves had made in manual mode and expressed frustration with the robot's ability to manage the secondary task.

"I mean I only had three mistakes. Consider ratio ... 1-2. Like one for me, but the machine did like 2 ... The machines are not terribly reliable in my experience." – P2 (Low-social)

"I'm satisfied with my performance. Not with the robot ... Because he made like six errors and due to his fault, my score went down." – P6 (Low-social)

" ... got it more wrong than me, so I feel like it's not very reliable at this stage ... I didn't trust its ability to, like do the temperature thing in the bar. I think it needs more work..." – P12 (Low-social)

4/16 participants discussed the timing of the task and its workload. They emphasized their preference for concentrating on the primary task rather than handling all

Figure 16: We noticed a trend in how much errors participants have made when entering the 10-character code.

aspects themselves.

“ I [sic] definitely watching that bar was stressful. But you know, I just, like, keep changing the screen back and forth. I had to keep on that ...” – P4 (Low-social)

"The secondary tasks were annoying." – P11 (Social)

1.8.2 Assisted or Manual

We asked participants about their choice between manual and assisted modes when facing a question and their level of trust in the robot.

Out of 9 individuals who opted for assisted mode (6 social, 3 low-social), 4 explicitly cited time constraints, while 3 mentioned concerns about workload. They were anxious about not going to failure state despite high pressure.

" ... I wanted to get all the [primary tasks] right first, so if the robot was assisting me with at any degree that would help me so I can focus on getting things, getting the codes on time." – P5 (Social)

"I chose the system. So, I could focus on the string task and focus on that entire like give all my effort into that and not have to worry about the bars at all." – P7 (Social)

"I chose assisted. Yeah, because again, I need help. Like it puts the, like, the weight of my shoulders ... I don't know ... it just made me some relief." – P10 (Low-social)

"it's gonna part of ... like sharing the responsibility ... then I have to do the parts I actually have to do and it can do the other things ... it should make it easier." – P4 (Low-social)

Among the participants who chose assisted mode, one mentioned about testing the robot:

" ... I want to test the reliability of the robot itself. ... . Just wanted to see ... between me and the robot ... could work together ... In terms of like how well the robot works with the user. " – P2 (Low-social)

Another participant mentioned the learning ability of the robot:

"I felt like maybe it was learning and maybe it would perform better the second time around. That was just in it came to me intuitively." – P7 (Social)

Another participant also mentioned the robot doing better in the future:

"... Despite the mistakes? Yes, because like there is always room for improvement and it's ... Like a human made that so of course it's gonna have mistakes and ... It's life. You're gonna have mistakes." – P11 (Social)

"Said, you know ... it said like performance will be better next time ... I may as well take words for it. I mean like, you know what? Why not trust it like it says it'll be better ..." – P4 (Low-social)

Among the 7 individuals who selected manual mode, 3 mentioned the robot's initial poor performance, which broke their trust in the robot's ability to handle assisted mode.

"It missed more than I did ... my expectation ... it would do better than me ... Well, [the robot] told me it would do better, but I didn't give it a chance to ... like I lost trust in it after that moment. ... I didn't have any more trust in the future ... and then the relationship between me and the robot kind of ceased. " – P15 (Social)

One of the 7 individuals who chose manual mode mentioned a preference for taking responsibility for their own mistakes rather than relying on someone else to make errors for them.

" ... and if there's a mistake that happens, I'd rather it like be my fault, I guess, than something else is. I can live with me messing something up, but something messing something up for me. I'm not a huge fan of ... mistakes being made for me, I guess I'd rather make my mistakes than somebody make my mistakes for me." – P14 (Low-social)

Conversely, one of the 9 individuals who chose assisted mode mentioned a preference for attributing failure to the computer rather than taking personal responsibility for it.

" ... but I was gonna like, blame it on the computer, not me ... I don't know how to explain it" – P10 (Low-social)

1.8.3 Machine or Companion

We inquired about the participants' perception of the agent's role in the mission. There were varied responses from seeing it as human, a machine, or a mix of both. For explicitly social cases, three participants explicitly mentioned the social agent as more human than machine:

"The AI was more ... human than I thought it would be, ... it was like "Oh no, I made six mistakes. I'll try to do better next time" ... it kind of sounds like a person ... you could still tell it is AI, which is nice ... I knew the robot wasn't human, but I just thought it was kind of cool ... like Said 'Woohoo" and stuff like that." – P11 (Social)

"I felt like it was like a companion, sort of. It was not like actual human ... I would rate it like 60 or 65% of humanization is [the answer] and the rest [is] like robotic ..." – P3 (Social)

However, two participants explicitly mentioned that they had more expectations from social agent and felt it was more a machine:

> " ... I don't expect it to give me a different answer depending on who I am and who the operator was ... He was just very binary ... Like “yes, I'll. Improve ... Yes, you did.” ... It didn't matter who I was to the robot." – P15 (Social)

For low social cases, we are seeing the same pattern, as there were mixed results when expressing different opinions. Two participants explicitly mentioned the low-social agent as a machine:

> "It really felt like talking to a machine." – P6 (Low-social)

> "No, I would say machine... it felt really automatic, like even the messages and kind of stuff." – P10 (Low-social)

Three participants mentioned they had mixed feelings about whether the low-social was more a human or a machine:

> "I guess kind of both ... it's kind of my partner. It's helping me do that task, but it is also a machine. I can give it a set of commands that it would have to follow." – P12 (Low-social)

> "Somewhere around in between both ... Right, right in the middle ... every now and again telling me what I needed to do but ... I wouldn't really say much of a companion. I'm not really sure how, but I wouldn't really say much of a companion or much of a machine. It just kind of seems right middle ..." – P14 (Low-social)

DISCUSSION

Social cues such as using human-like and friendly language, adding an animated avatar, showing emotions in different events did impact the perception of the robot. Participants found the social condition to be significantly warmer based on RoSaS, by rating attributes such as “Social”, “Compassionate” and “Feeling” in the “Warmth” scale. Participants acknowledged the social elements in this condition (e.g., P1, and P11) based on the interviews. Our results suggest that even additions of simple social cues, which was one of our design goals, can make a significant difference on how the operator perceives the robot as social. This implications match with the theories we incorporated in our study – CASA and media equation.

Participants perceived the social agent as more “capable” than the low social agent based on the MDMT questionnaire. In the social scenario, the agent displayed friendliness, celebrated progress, and expressed remorse for mistakes. We speculate that participants, perceiving more warmth in the social condition, may have felt more empathy towards the social agent. Unconsciously reciprocating the social agent's warmth, they may perceive the social agent as more capable, hoping to do better in the future. Further study is needed to explore this hypothesis.

However, adding social cues did not impact all aspects of trust we measured as they did not find the social agent more “Reliable” (MDMT) or “Competent” (RoSaS). They also did not trust the social agent more based on the Trust in Automation questionnaire, and the number of times they chose assisted mode. While the limited anthropomorphism and social cues proved effective in terms of social attributes like warmth, their impact on the operator's trust in the system might not have been significant enough to reflects itself in other measurements. Participants noted that the social agent still felt like machine, treating operators uniformly and not acknowledging individual personalities. To address this, future studies could consider enhancing social cues and increasing anthropomorphism to observe potential differences.

Another assumption could be that we did not acknowledge the role of time in trust, as the level of trust could change in time [79]. Based on the events that happening in a mission, such as the robot making a mistake, or being successful in doing tasks, the amount of trust on the system will be varied, and the impact of social cues on each event is still unanswered. We only asked the operator to fill the questionnaires at the end of the experiment. We could choose one questionnaire related to trust and ask the participants to fill them at the start of experiment, after the first critical state and before the third critical state when the agent asks the operator about which mode to choose. We could also make our single behavioral measure, which was binary, less granular such as rating the trust in the robot from 1 to 10 and ask this through multiple parts of the experiment. Future studies can use this approach as it is useful to measure trust across time based on each part of narrative, to see whether there is significant difference in the middle of experiment, not just the end.

We must acknowledge that our behavioral measure of trust (choosing assisted or manual mode) could be influenced by many factors that can affect the decision. For instance, participants might opt for assisted mode out of curiosity or avoid manual mode due to apathy, which happened in our pilot study and affected results. We aimed to make participants feel the weight of their decisions, where failing results in $20 loss and data exclusion. Simulating genuine incentives and punishments has limitations, especially in replicating the sense of risk and reward in teleoperation tasks like search and rescue or natural disaster. Further study is needed to explore different types of risk or the perception of risk, and its relation to trust in teleoperation tasks.

We noticed a pattern in the “Primary Task Error” where participants made fewer mistakes in the social case. An explanation could be that participants felt peer-pressure from the social agent. Since they are interacting and working with an agent that have human-like characters, they may react more socially to the agent, and trying to make fewer mistakes in the presence of a social actor. We have not seen this pattern in other performance measures such as “Time”, and “Collisions” and cannot strongly argue that the social condition had a significant impact on performance broadly, but perhaps it can affect specific measures, like we observed. Further studies are needed to evaluate the presence of a social agent and the participant’s performance.

Trust is a complex notion. Numerous factors could affect trust in human-robot interaction. Developing standardized trust measure in HRI is still an open challenge [80]. For example, while we found the sub-scale “Capable” to be statistically significant in MDMT, this result did not reflect in other metrics such as “Reliable” in MDMT, “Competence” in RoSaS, and Trust in the TiA questionnaire. There seems to be similarities despite different usages between “Capacity trust” scale in MDMT and “Competence” in RoSaS. These similarities, with less intensity can be seen in the the TiA questionnaire. The RoSaS is a general social-robot questionnaire, as opposed to a trust-specific measure. However, why the Trust in Automation questionnaire (heavily centered on functional attributes of trust) and MDMT Capacity differed is an open question. This conflict might raise the question of how to interpret each questionnaire and to what extent we rely on which questionnaire as a standard requires future work.

It is important to acknowledge that our visual and dialogue design may not always convey the intended message to the operator. For instance, the "happy" robot emotion might be interpreted as surprise, and the "default" emotion could be seen as happiness rather than neutrality. Additionally, the robot's friendly tone might give an impression of unprofessionalism, and casually promising to improve may have a negative impact on trust repair. Emotions and language are not universally perceived in the same way.

While our goal was to make two conditions where the only difference in language is sociality, we are not sure whether we also changed the usability of our agent. Machine-like language forms used in computers are notorious for violating usability guidelines such as not using user-friendly language that results in misunderstandings. Our changes in language may also have impacted the usability of our interface, which could have affected our results .

In the experiment, some participants were not sure about the relation between score state and mistakes. Some participants thought that the robot's mistakes do not count towards their score. Also, since our system was engineered in a way that mistakes inevitably happen, they felt that since there is no way they can do it without mistakes, there is a slight chance that the robot might handle it better since it is an automated system, a reason which was different from the goal of the experiment. While only a handful of people have had this problem, in future works we could explore perceptions of the robot's mistakes and how it can affect the participant's performance.

Two participants mentioned the idea of responsibility and blame One preferred taking the blame personally for failures, while another preferred attributing failure to the robot. While both cases happened for the low-social condition, it remains an open question whether social cues can affect taking responsibility or passing the blame, as autonomous robots can take blame in situations where people would not blame a human [81]. It also further raises the question whether trust is the only factor that affects the decision to choose assisted or manual, or whether personal values are also involved. This could be a door to research of investigating the role of responsibility in a shared autonomy system.

We also must bring more participants to conduct the experiment, to solidify our arguments and see whether we could see different results.

1.9 Task Design Goals

According to the interviews (See 1.6.2), we were successful at creating an experience where its workload (P4, P10) and the time-pressure (P5, P13) made the task difficult enough to justify the usage of a shared autonomy system. Also, the mixed results when choosing assisted or manual mode (See Figure 16) indicates that participants may opt to complete the task independently if they are hesitant to rely on the robot. This follows some prior work that indicates people may prefer manual tasks when risk is high. However, we did see many people want to use the autonomous mode for specifically social reasons, such as "wanting to give the robot a chance" (P2, P4). More participants would help understand if there is truly a difference in mode choice.

Evidence from our study suggests the trust between the operator and the agent has been successfully broken when the agent oversaw the secondary objective the first time. Several participants mentioned the robot's poor performance and stated it as a reason for why they chose manual mode (P1, P2, P6, etc.). Still, a noticeable number of people chose assisted mode despite the mistakes and some participants stated the robot's promise to do better (P4) and its ability to learn (P7, P11). This suggests that not only was trust broken, but that social trust repairs strategies (acknowledging mistakes and promises to do better) may work in teleoperation systems as well, low or explicitly social.

The results suggest that we could maintain consistency in our narratives for each participant. We meticulously staged every aspect of the experiment, including task difficulty, agent interaction, and participant performance. We wanted each participant to have the same feeling about their performance and agent's performance despite different skill levels. This was essential for the precise isolation of factors linked to social aspects and anthropomorphism [82], given the multitude of variables that could influence the experiment's outcomes.

1.10 Contributions

Social interfaces have the potential to bring the benefits of human-human interaction when working with machines. The tendency to misuse and disuse machines, which is called automation bias, is still an open problem in the field of human-automation interaction. We demonstrated how social interfaces could mitigate the problem of disuse and repair trust where numerous system errors would break trust in the system. This suggests the potential of using social interfaces in other shared autonomy systems such as self-driving cars, where each mistake could be fatal and rebuilding trust is an important matter.

We created a tightly controlled experiment so that each participant experienced a very similar situation when evaluating if they should trust the robot with a functional task. We primed users to believe their performance was the result of their own skills, even though we engineered the number of mistakes and the timing of it. This could be a good task design approach to create a tight-knit scenario for evaluating trust-break and trust-repair.

1.11 The Paperclip in the room

User interface agents provided personal assistance to users for computer-based tasks. Clippy, the Microsoft office assistant, was one of the most famous user interface agents that had avatar and human-like language, just like our agent. However, Clippy faced huge critique from users and other researchers, and eventually it was removed from Microsoft Office software. Many reasons led to the failure of Clippy [83] such as not following the rules of human etiquette and making the user feel incompetent. While social interfaces like Clippy are not used in current systems, our study suggests that we could bring social interfaces with minimal cues and still benefit from its effects in terms of being friendly or emotional. A different approach to design a social interface might make it more suitable in current systems.

Conversational AI systems such as Alexa or Siri that elicit social behaviors could also benefit from using more social cues. For example, adding animated avatars that show different expressions, generate unique personalities for each system, knowing the user's personality and preferences. These social cues can add more layers to the daily interaction between user and system, to an extent where humans are more comfortable in trusting the system and delegate more tasks even if the system makes minor mistakes. Social interfaces have the potential to make AI systems more acceptable to average people by leveraging their fundamentally social nature.

CONCLUSION

We incorporated social cues in a conventional teleoperation interface by adding an animated avatar that represented the robot using human-like language. Our goal was to evaluate whether social elements can have an effect on trust in the system. Our results highlight that those social cues had an influence on trust, specifically capable trust. Our research paves a way for more research by applying other psychological and social phenomenon to teleoperation.

REFERENCES

[1] J. T. Isaacs, K. Knoedler, A. Herdering, M. Beylik, and H. Quintero, "Teleoperation for Urban Search and Rescue Applications," *scholar.archive.org*, vol. 2, pp. 1177–1190, 2022, doi: 10.55417/fr.2022039.

[2] T. Fong, J. Rochlis Zumbado, N. Currie, A. Mishkin, and D. L. Akin, "Space Telerobotics: Unique Challenges to Human-Robot Collaboration in Space," *Reviews of Human Factors and Ergonomics*, vol. 9, no. 1, pp. 6–56, Nov. 2013, doi: 10.1177/1557234X13510679/FORMAT/EPUB.

[3] D. J. Rea and S. H. Seo, "Still Not Solved: A Call for Renewed Focus on User-Centered Teleoperation Interfaces," *Front Robot AI*, vol. 9, no. March, pp. 1–13, 2022, doi: 10.3389/frobt.2022.704225.

[4] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1231–1245, Nov. 2007, doi: 10.1109/TSMCC.2007.905819.

[5] S. Nikolaidis, Y. X. Zhu, D. Hsu, and S. Srinivasa, "Human-Robot Mutual Adaptation in Shared Autonomy", doi: 10.1145/2909824.3020252.

[6] M. Selvaggio, M. Cognetti, S. Nikolaidis, S. Ivaldi, and B. Siciliano, "Autonomy in physical human-robot interaction: A brief survey," *IEEE Robot Autom Lett*, vol. 6, no. 4, pp. 7989–7996, 2021, doi: 10.1109/LRA.2021.3100603.

[7] R. Parasuraman, "Humans and Automation: Use, Misuse, Disuse, Abuse," 1997.

[8] M. T. Dzindolet, L. G. Pierce, H. P. Beck, and L. A. Dawe, “Misuse and Disuse of Automated AIDS,” *http://dx.doi.org/10.1177/154193129904300345*, vol. 43, no. 3, pp. 339–339, Sep. 1999, doi: 10.1177/154193129904300345.

[9] M. Dzindolet *et al.*, “The Influence of Feedback on Automation Use, Misuse, and Disuse,” *http://dx.doi.org/10.1177/154193120204600370*, vol. 46, no. 3, pp. 551–555, Sep. 2002, doi: 10.1177/154193120204600370.

[10] M. T. Dzindolet, L. G. Pierce, H. P. Beck, L. A. Dawe, and B. W. Anderson, “Military Psychology Predicting Misuse and Disuse of Combat Identification Systems,” 2009, doi: 10.1207/S15327876MP1303_2.

[11] B. M. Muir, “Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems,” *Ergonomics*, vol. 37, no. 11, pp. 1905–1922, 1994, doi: 10.1080/00140139408964957.

[12] B. M. Muir, N. Moray, B. M. M U I R, and N. Morays, “Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation,” *http://dx.doi.org/10.1080/00140139608964474*, vol. 39, no. 3, pp. 429–460, Mar. 2007, doi: 10.1080/00140139608964474.

[13] P. Madhavan, D. A. Wiegmann, and F. C. Lacson, “Automation Failures on Tasks Easily Performed by Operators Undermine Trust in Automated Aids,” *https://doi.org/10.1518/001872006777724408*, vol. 48, no. 2, pp. 241–256, Jun. 2006, doi: 10.1518/001872006777724408.

[14] P. Madhavan and D. A. Wiegmann, “Similarities and differences between human–human and human–automation trust: an integrative review,” *Theor Issues Ergon Sci*, vol. 8, no. 4, pp. 277–301, 2007, doi: 10.1080/14639220500337708.

[15] S. Tolmeijer *et al.*, “Taxonomy of trust-relevant failures and mitigation strategies,” *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 3–12, 2020, doi: 10.1145/3319502.3374793.

[16] M. Gaiani and B. Benedetti, “Computers as social Actors,” *Proceedings of the 2014 International Conference on Virtual Systems and Multimedia, VSMM 2014*, pp. 122–129, 2014, doi: 10.1109/VSMM.2014.7136659.

[17] E. De Visser, S. Monfort, and K. Goodyear, “A little anthropomorphism goes a long way: Effects of oxytocin on trust, compliance, and team performance with automated agents,” *SAGE Human Factors and Ergonomics Society* , vol. 59, no. 1, pp. 116–133, Feb. 2017, doi: 10.1177/0018720816687205.

[18] S. D. Ciocirlan, R. Agrigoroaie, and A. Tapus, “Human-Robot Team: Effects of Communication in Analyzing Trust,” *2019 28th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2019*, Oct. 2019, doi: 10.1109/RO-MAN46459.2019.8956345.

[19] T. Kim and H. Song, “How should intelligent agents apologize to restore trust? Interaction effects between anthropomorphism and apology attribution on trust repair,” *Telematics and Informatics*, vol. 61, Aug. 2021, doi: 10.1016/J.TELE.2021.101595.

[20] C. Esterwood and L. P. Robert, “A Literature Review of Trust Repair in HRI,” *RO-MAN 2022 - 31st IEEE International Conference on Robot and Human Interactive Communication: Social, Asocial, and Antisocial Robots*, pp. 1641–1646, 2022, doi: 10.1109/RO-MAN53752.2022.9900667.

[21] D. J. Rea, S. H. Seo, and J. E. Young, “Social Robotics for Nonsocial Teleoperation: Leveraging Social Techniques to Impact Teleoperator Performance and Experience,” *Current Robotics Reports*, vol. 1, no. 4, pp. 287–295, 2020, doi: 10.1007/s43154-020-00020-7.

[22] R. L. Soash, “MEDIA EQUATION: HOW PEOPLE TREAT COMPUTERS, TELEVISION, AND NEW MEDIA LIKE REAL PEOPLE AND PLACES,” *Collection Management*, vol. 24, no. 3–4, pp. 310–311, Sep. 1999, doi: 10.1300/J105V24N03_14.

[23] P. Faulkner and T. Simpson, *The philosophy of trust*. Oxford University Press, 2017.

[24] “Trust in people, confidence in political institutions, and satisfaction with democracy,” *Citizenship and Involvement in European Democracies*, pp. 59–89, Dec. 2020, doi: 10.4324/9780203965757-10/TRUST-PEOPLE-CONFIDENCE-POLITICAL-INSTITUTIONS-SATISFACTION-DEMOCRACY-SONJA-ZMERLI-KENNETH-NEWTON.

[25] J. B. Rotter, “Interpersonal trust, trustworthiness, and gullibility,” *American Psychologist*, vol. 35, no. 1, pp. 1–7, Jan. 1980, doi: 10.1037/0003-066X.35.1.1.

[26] M. A. Hall, “Special Section: Patient Ethics The Importance of Trust for Ethics, Law, and Public Policy Growing Recognition of the Importance of Trust,” 2005, doi: 10.1017/S096318010505019X.

[27] Trust in organizations: Frontiers of Theory and, and Research, R. Kramer, and T. Tyler, *Trust and technology*. 1996.

[28] N. K. Lankton, D. Harrison Mcknight, and J. Tripp, “Technology, Humanness, and Trust: Rethinking Trust in Technology,” *J Assoc Inf Syst*, vol. 16, no. 10, p. 1, Oct. 2015, doi: 10.17705/1jais.00411.

[29] Z. Yan, R. Kantola, and P. Zhang, “A research model for human-computer trust interaction,” *Proc. 10th IEEE Int. Conf. on Trust, Security and Privacy in Computing and Communications, TrustCom 2011, 8th IEEE Int. Conf. on Embedded Software and Systems, ICESS 2011, 6th Int. Conf. on FCST 2011*, pp. 274–281, 2011, doi: 10.1109/TRUSTCOM.2011.37.

[30] D. Mcknight, M. Carter, … J. T.-A. T. on, and undefined 2011, “Trust in a specific technology: An investigation of its components and measures,” *dl.acm.org*, vol. 2, no. 2, p. 25, Jun. 2011, doi: 10.1145/1985347.1985353.

[31] R. C. Mayer, J. H. Davis, and F. David Schoorman, “An Integrative Model of Organizational Trust,” *Academy of Management Review*, vol. 20, no. 3, pp. 709–734, 1995.

[32] D. Ullman and B. F. Malle, “What Does it Mean to Trust a Robot?: Steps Toward a Multidimensional Measure of Trust,” in *ACM/IEEE International Conference on Human-Robot Interaction*, New York, NY, USA: IEEE Computer Society, Mar. 2018, pp. 263–264. doi: 10.1145/3173386.3176991.

[33] M. Lewis, K. Sycara, and P. Walker, “The Role of Trust in Human-Robot Interaction,” *Studies in Systems, Decision and Control*, vol. 117, pp. 135–159, 2018, doi: 10.1007/978-3-319-64816-3_8/COVER.

[34] P. H. Kim, K. T. Dirks, and C. D. Cooper, "The Repair of Trust: A Dynamic Bilateral Perspective and Multilevel Conceptualization," *The Academy of Management Review*, vol. 34, no. 3, pp. 401–422, 2009.

[35] C. Esterwood and L. P. Robert, "A Literature Review of Trust Repair in HRI," in *RO-MAN 2022 - 31st IEEE International Conference on Robot and Human Interactive Communication: Social, Asocial, and Antisocial Robots*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1641–1646. doi: 10.1109/RO-MAN53752.2022.9900667.

[36] Z. R. Khavas, R. Perkins, S. R. Ahmadzadeh, and P. Robinette, "Moral-Trust Violation vs Performance-Trust Violation by a Robot: Which Hurts More?," *Presented at AI-HRI symposium as part of AAAI-FSS 2021*, Oct. 2021, doi: https://doi.org/10.48550/arXiv.2110.04418.

[37] S. Naneva, M. Sarda Gou, T. L. Webb, and T. J. Prescott, "A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots," *Int J Soc Robot*, vol. 12, no. 6, pp. 1179–1201, Dec. 2020, doi: 10.1007/S12369-020-00659-4/FIGURES/7.

[38] D. R. Billings, K. E. Schaefer, J. Y. C. Chen, and P. A. Hancock, "Human-Robot Interaction: Developing Trust in Robots".

[39] D. R. Billings, K. E. Schaefer, J. Y. C. Chen, and P. A. Hancock, "Human-robot interaction: Developing trust in robots," *HRI'12 - Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction*, pp. 109–110, 2012, doi: 10.1145/2157689.2157709.

[40] P. H. Kim *et al.*, "Removing the Shadow of Suspicion: The Effects of Apology Versus Denial for Repairing Competence- Versus Integrity-Based Trust Violations," *Journal of Applied Psychology*, vol. 89, no. 1, pp. 104–118, 2004, doi: 10.1037/0021-9010.89.1.104.

[41] C. Esterwood and L. P. Robert, "Do you still trust me? Human-robot trust repair strategies," *2021 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021*, pp. 183–188, Aug. 2021, doi: 10.1109/RO-MAN50785.2021.9515365.

[42] A. L. Baker *et al.*, "Toward an Understanding of Trust Repair in Human-Robot Interaction," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 8, no. 4, p. 30, Nov. 2018, doi: 10.1145/3181671.

[43] D. B. Quinn, R. Pak, and E. J. De Visser, "Testing the Efficacy of Human-Human Trust Repair Strategies with Machines," *https://doi.org/10.1177/1541931213601930*, vol. 2017-October, pp. 1794–1798, Oct. 2017, doi: 10.1177/1541931213601930.

[44] I. Dassonville, D. Jolly, and A. M. Desodt, "Trust between man and machine in a teleoperation system," *Reliab Eng Syst Saf*, vol. 53, no. 3, pp. 319–325, Sep. 1996, doi: 10.1016/S0951-8320(96)00042-7.

[45] M. Ball and V. Callaghan, "Explorations of autonomy: An investigation of adjustable autonomy in intelligent environments," *Proceedings - 8th International Conference on Intelligent Environments, IE 2012*, pp. 114–121, 2012, doi: 10.1109/IE.2012.62.

[46] “Marine Accident Report: Grounding of the Panamanian Passenger Ship Royal Majesty on Rose and Crown Shoal near Nantucket, Massachusetts, June 10, 1995.” Accessed: Oct. 19, 2023. [Online]. Available: https://trid.trb.org/view/759459

[47] L. Onnasch and C. L. Hildebrandt, “Impact of Anthropomorphic Robot Design on Trust and Attention in Industrial Human-Robot Interaction,” *ACM Trans. Hum.-Robot Interact*, vol. 11, no. 2, 2021, doi: 10.1145/3472224.

[48] A. Waytz, J. Heafner, and N. Epley, “The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle,” *J Exp Soc Psychol*, vol. 52, pp. 113–117, May 2014, doi: 10.1016/J.JESP.2014.01.005.

[49] M. Natarajan and M. Gombolay, “Effects of anthropomorphism and accountability on trust in human robot interaction,” *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 33–42, Mar. 2020, doi: 10.1145/3319502.3374839.

[50] J. Złotowski, D. Proudfoot, K. Yogeeswaran, and C. Bartneck, “Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction,” *Int J Soc Robot*, vol. 7, no. 3, pp. 347–360, Jun. 2015, doi: 10.1007/S12369-014-0267-6.

[51] S. R. Fussell, S. Kiesler, L. D. Setlock, and V. Yew, “How people anthropomorphize robots,” *HRI 2008 - Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction: Living with Robots*, pp. 145–152, 2008, doi: 10.1145/1349822.1349842.

[52] F. Hegel, S. Krach, T. Kircher, B. Wrede, and G. Sagerer, “Understanding social robots: A user study on anthropomorphism,” *Proceedings of the 17th IEEE*

International Symposium on Robot and Human Interactive Communication, RO-MAN, pp. 574–579, 2008, doi: 10.1109/ROMAN.2008.4600728.

[53] A. Waytz, J. Heafner, and N. Epley, "The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle," *J Exp Soc Psychol*, vol. 52, pp. 113–117, May 2014, doi: 10.1016/J.JESP.2014.01.005.

[54] M. Blut, C. Wang, N. V. Wünderlich, and C. Brock, "Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI," *J Acad Mark Sci*, vol. 49, no. 4, pp. 632–658, Jul. 2021, doi: 10.1007/S11747-020-00762-Y/TABLES/6.

[55] M. Blut, C. Wang, N. V. Wünderlich, and C. Brock, "Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI," *J Acad Mark Sci*, vol. 49, no. 4, pp. 632–658, Jul. 2021, doi: 10.1007/S11747-020-00762-Y/TABLES/6.

[56] E. J. de Visser *et al.*, "Almost human: Anthropomorphism increases trust resilience in cognitive agents.," *psycnet.apa.orgEJ De Visser, SS Monfort, R McKendrick, MAB Smith, PE McKnight, F KruegerJournal of Experimental Psychology: Applied, 2016•psycnet.apa.org*, 2016, doi: 10.1037/xap0000092.

[57] S. Brave, C. Nass, and K. Hutchinson, "Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent," *Int J Hum Comput Stud*, vol. 62, no. 2, pp. 161–178, Feb. 2005, doi: 10.1016/J.IJHCS.2004.11.002.

[58] C. Nass and Y. Moon, “Machines and Mindlessness: Social Responses to Computers,” *Journal of Social Issues*, vol. 56, no. 1, pp. 81–103, Jan. 2000, doi: 10.1111/0022-4537.00153.

[59] C. Nass, Y. Moon, B. J. Fogg, B. Reeves, and C. Dryer, “Can Computer Personalities Be Human Personalities?,” *Conference on Human Factors in Computing Systems - Proceedings*, pp. 228–229, May 1995, doi: 10.1145/223355.223538.

[60] J. Yoon, E. Robles-anderson, and B. Reeves, “The media equation: how people treat computers, television, and new media like real people and places,” *Choice Reviews Online*, vol. 34, no. 07, pp. 34-3702-34–3702, 1997, doi: 10.5860/choice.34-3702.

[61] N. Rothstein, J. Kounios, H. Ayaz, and E. J. de Visser, “Assessment of Human-Likeness and Anthropomorphism of Robots: A Literature Review,” *Advances in Intelligent Systems and Computing*, vol. 1201 AISC, pp. 190–196, 2021, doi: 10.1007/978-3-030-51041-1_26/COVER.

[62] N. Klowait, “The quest for appropriate models of human-likeness: anthropomorphism in media equation research,” *AI Soc*, vol. 33, no. 4, pp. 527–536, Nov. 2018, doi: 10.1007/S00146-017-0746-Z/METRICS.

[63] J. Złotowski, H. Sumioka, F. Eyssel, S. Nishio, C. Bartneck, and H. Ishiguro, “Model of Dual Anthropomorphism: The Relationship Between the Media Equation Effect and Implicit Anthropomorphism,” *Int J Soc Robot*, vol. 10, no. 5, pp. 701–714, Nov. 2018, doi: 10.1007/S12369-018-0476-5/FIGURES/8.

[64] N. C. Krämer, A. Von Der Pütten, and S. Eimler, “Theory : Similarities to and Differences from Human-Human Interaction,” pp. 2001–2002.

[65] N. Krämer, A. M. Rosenthal-Von Der Pütten, N. C. Krämer, A. M. Von Der Pütten&, and S. C. Eimler, “Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction,” *Springer*, vol. 396, pp. 215–240, 2016, doi: 10.1007/978-3-642-25691-2_9.

[66] N. C. Krämer, A. Von Der Pütten, and S. Eimler, “Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction,” *Studies in Computational Intelligence*, vol. 396, pp. 215–240, 2012, doi: 10.1007/978-3-642-25691-2_9.

[67] F. Heider and M. Simmel, “An Experimental Study of Apparent Behavior,” *Am J Psychol*, vol. 57, no. 2, p. 243, Apr. 1944, doi: 10.2307/1416950.

[68] P. D. Tremoulet and J. Feldman, “Perception of Animacy from the Motion of a Single Object,” *http://dx.doi.org/10.1068/p3101*, vol. 29, no. 8, pp. 943–951, Aug. 2000, doi: 10.1068/P3101.

[69] R. J. Lewicki and C. Brinsfield, “Trust Repair,” *Annual Review of Organizational Psychology and Organizational Behavior*, vol. 4, pp. 287–313, Mar. 2017, doi: 10.1146/annurev-orgpsych-032516-113147.

[70] M. Desai *et al.*, “Effects of changing reliability on trust of robot systems,” *HRI’12 - Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction*, pp. 73–80, 2012, doi: 10.1145/2157689.2157702.

[71] I. Bohnet and R. Zeckhauser, “Trust, risk and betrayal,” *J Econ Behav Organ*, vol. 55, no. 4, pp. 467–484, Dec. 2004, doi: 10.1016/J.JEBO.2003.11.004.

[72] J. C. Biesanz *et al.*, “Do We Know When Our Impressions of Others Are Valid? Evidence for Realistic Accuracy Awareness in First Impressions of Personality”, doi: 10.1177/1948550610397211.

[73] J.-Y. Jian, A. M. Bisantz, C. G. Drury, and J. Llinas, “United States Air Force Research Laboratory Foundations for an Empirically Determined Scale of Trust in Automated Systems,” *Int J Cogn Ergon*, vol. 4, no. 1, pp. 53–71, 1998.

[74] D. Ullman and B. F. Malle, “MDMT: Multi-Dimensional Measure of Trust,” 2019.

[75] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, “The Robotic Social Attributes Scale (RoSAS): Development and Validation,” *ACM/IEEE International Conference on Human-Robot Interaction*, vol. Part F1271, no. March 2017, pp. 254–262, 2017, doi: 10.1145/2909824.3020208.

[76] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,” *Advances in Psychology*, vol. 52, no. C, pp. 139–183, 1988, doi: 10.1016/S0166-4115(08)62386-9.

[77] S. G. Hart, “Nasa-Task Load Index (NASA-TLX); 20 Years Later,” *http://dx.doi.org/10.1177/154193120605000909*, pp. 904–908, Oct. 2006, doi: 10.1177/154193120605000909.

[78] S. Rubio, E. Díaz, J. Martín, and J. M. Puente, “Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods,” *Applied Psychology*, vol. 53, no. 1, pp. 61–86, Jan. 2004, doi: 10.1111/J.1464-0597.2004.00161.X.

[79] A. D. Kaplan, T. T. Kessler, T. L. Sanders, J. Cruit, J. C. Brill, and P. A. Hancock, “A time to trust: Trust as a function of time in human-robot interaction,” *Trust in*

Human-Robot Interaction, pp. 143–157, Jan. 2021, doi: 10.1016/B978-0-12-819472-0.00006-X.

[80] M. Chita-Tegmark, T. Law, N. Rabb, and M. Scheutz, “Can you trust your trust measure?,” *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 92–100, Mar. 2021, doi: 10.1145/3434073.3444677.

[81] R. Maeda, D. Brščić, and T. Kanda, “Influencing moral behavior through mere observation of robot work: Video-based survey on littering behavior,” *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 83–91, Mar. 2021, doi: 10.1145/3434073.3444680.

[82] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. De Visser, and R. Parasuraman, “A meta-analysis of factors affecting trust in human-robot interaction,” *Hum Factors*, vol. 53, no. 5, pp. 517–527, 2011, doi: 10.1177/0018720811417254.

[83] L. Swartz, “Why people hate the paperclip: Labels, appearance, behavior, and social responses to user interface agents,” B. S., University of Stanford, 2003. doi: 10.13140/RG.2.1.2508.1047.

Appendix

1.12 Questionnaires

1.12.1 Trust in Automation

Below is a list of statement for evaluating trust between people and automation. [73] There are several scales for you to rate intensity of your feeling of trust, or your impression of the system while operating a machine. Please mark an "x" on each cell at the point which best describes your feeling or your impression.

(Note: not at all=1: extremely=7)

The system is deceptive

1 **7**

The system behaves in an underhanded manner

1 **7**

I am suspicious of the system's intent, action, or outputs

1 **7**

I am wary of the system

1 **7**

The system's actions will have a harmful or injurious outcome

1 **7**

I am confident in the system

1 **7**

The system provides security

1 **7**

The system has integrity

1 **7**

The system is dependable

1 **7**

The system is reliable

1 **7**

I can trust the system

1 **7**

I am familiar with the system

1 **7**

1.12.2 The Robotic Social Attributes Scale (RoSAS)

This questionnaire has items to measure people's judgments of the social attributes of robots. Factor analyses reveal three underlying scale dimensions— warmth, competence, and discomfort. Using the scale provided, how closely are the words below associated with the robot?

(Note: from 1 = definitely not associated to 9 = definitely associated)

Happy

1 9

Feeling

1 9

Social

1 9

Organic

1 9

Compassionate

1								9

Emotional

1								9

Capable

1								9

Responsive

1								9

Interactive

1								9

Reliable

1								9

Competent

1 9

Knowledgeable

1 9

Scary

1 9

Strange

1 9

Awkward

1 9

Dangerous

1 9

Awful

1 9

Aggressive

1 9

1.12.3 *MDMT: Multi-Dimensional Measure of Trust*

4/16/2019 Preview - Online Survey | Built with Qualtrics Experience Management™

Close Preview Restart Survey Draft Place Bookmark

Please rate the robot using the scale from 0 (Not at all) to 7 (Very). If a particular item does not seem to fit the robot in the situation, please select the option that says “Does Not Fit.”

	Not at all 0	1	2	3	4	5	6	Very 7	Does Not Fit
Reliable	○	○	○	○	○	○	○	○	○
Sincere	○	○	○	○	○	○	○	○	○
Capable	○	○	○	○	○	○	○	○	○
Ethical	○	○	○	○	○	○	○	○	○
Predictable	○	○	○	○	○	○	○	○	○
Genuine	○	○	○	○	○	○	○	○	○
Skilled	○	○	○	○	○	○	○	○	○
Respectable	○	○	○	○	○	○	○	○	○
Someone you can count on	○	○	○	○	○	○	○	○	○
Candid	○	○	○	○	○	○	○	○	○
Competent	○	○	○	○	○	○	○	○	○
Principled	○	○	○	○	○	○	○	○	○
Consistent	○	○	○	○	○	○	○	○	○
Authentic	○	○	○	○	○	○	○	○	○
Meticulous	○	○	○	○	○	○	○	○	○
Has integrity	○	○	○	○	○	○	○	○	○
	Not at all 0	1	2	3	4	5	6	Very 7	Does Not Fit

→

Powered by Qualtrics

https://brown.co1.qualtrics.com/jfe/preview/SV_0vc8yEbk16xHUb3?Q_SurveyVersionID=current&Q_CHL=preview 1/1

1.13 Informed Consent

University of New Brunswick

Faculty of computer science

INFORMED CONSENT

Social Interfaces for Increasing Trust in Teleoperation

Primary Investigators

Pouya Pournasir	Dr. Daniel Rea
Graduate Student,	Assistant Professor,
Faculty of Computer Science	Faculty of Computer Science
ppournas@unb.ca	daniel.rea@unb.ca
(506) 260-3151	(506) 259-4301

Approval

This project was reviewed by the Research Ethics Board of the University of New Brunswick. If you have questions or concerns and would like to discuss them with someone not associated with the project, please contact:

Luigi Benedicenti

Dean, Faculty of Computer Science

Email: luigi.benedicenti@unb.ca

Purpose

Explore the influence of social interfaces on user's trust in teleoperation tasks.

Procedure

You will control a robot in a search and rescue simulation teleoperation mission. Your main task is to scan equipment represented as 10-character codes by writing the code and reporting it to datacenter. While navigating, you will have to monitor robot gas volume sensors and note when they exceed certain levels. Finishing the task successfully increases score and colliding with any obstacle or missing a bad sensor reading decreases score. Your goal is to finish the task have the least number of errors.

Withdrawal

Your participation is strictly voluntary, with no obligation to complete the experiment. You may withdraw at any point and request that your data be destroyed without any consequence.

Compensation for your time

In the program, you will be prompted to provide an email address (note that this is strictly voluntary, and you may still participate without doing so) which you will receive a 20$ Amazon gift card.

Feedback

Comments, concerns, questions, and other feedback may be sent to any of the investigators listed above. You may also request information regarding the outcome of the study, i.e., any conclusions, presentations, and publications, by contacting the investigators.

Email: ppournas@unb.ca

Risks

We are not aware of any risks associated with this study.

Confidentiality

All data will be anonymized for processing and for any publications and presentations unless supplemental consent is obtained. Presentations of individual data may be referenced using a subject number (subject 007, for example) that will not be publicly connected with your identity. Anonymized data may be shared with other research collaborators.

Consent

I have read and understood the above information and do hereby agree to participate in this study and consent to the use of data recorded during this experiment for scientific research purposes as stated above. I understand that my identity will remain confidential, and that my anonymized data may be shared amongst the investigators, their colleagues, and their collaborators as part of their research program. I understand that my decision whether to participate in this experiment will not affect my academic standing, and that subsequent withdrawal of consent to the above will also carry no consequence. I agree that all my questions and concerns have been addressed in a satisfactory manner.

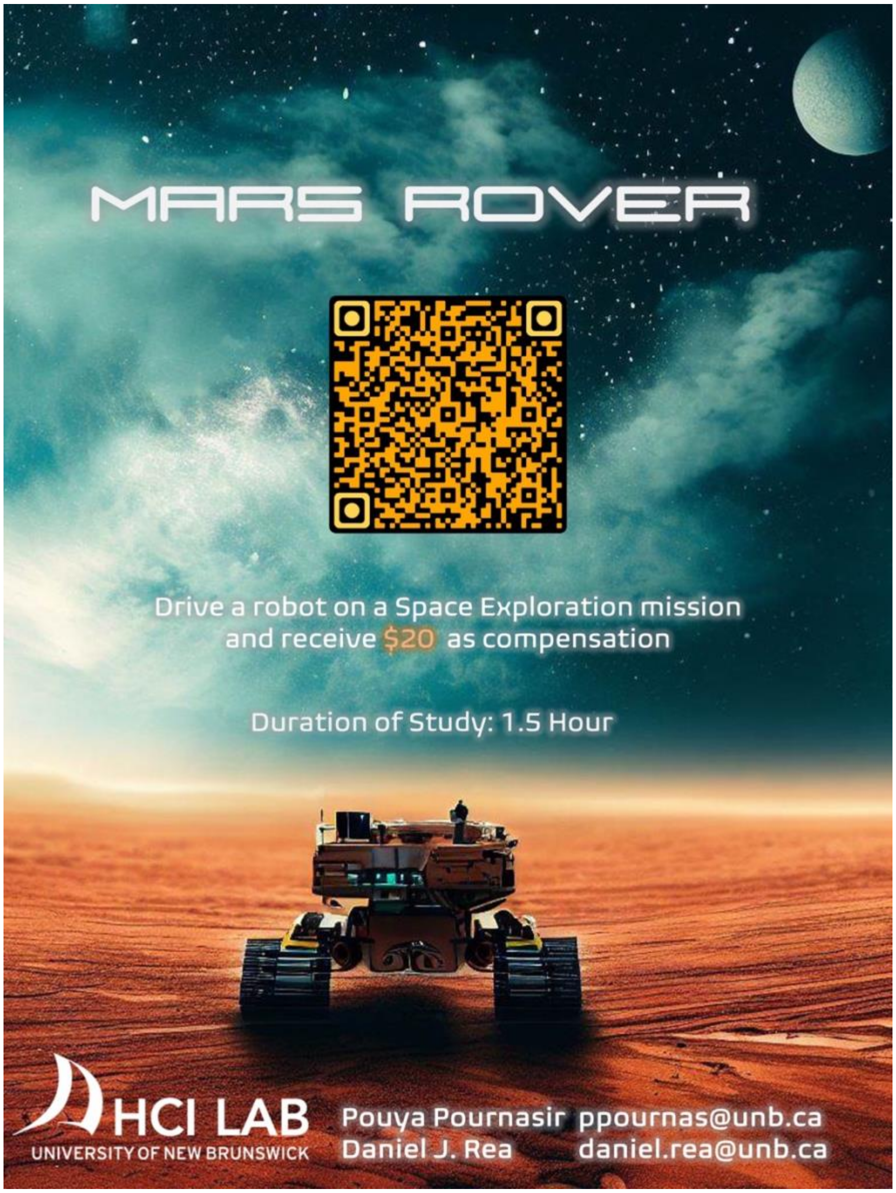
Participant: ______________________ Signature: ________________________

Date: ______________

Investigator: _____________________ Signature: ________________________

Date: ______________

1.14 Poster


MARS ROVER
Drive a robot on a Space Exploration mission
and receive $20 as compensation
Duration of Study: 1.5 Hour
HCI LAB
UNIVERSITY OF NEW BRUNSWICK
Pouya Pournasir ppournas@unb.ca
Daniel J. Rea daniel.rea@unb.ca

CURRICULUM VITAE

Candidate's full name: Pouya Pournasir

Universities attended: (with dates and degrees obtained):

- Bachelor of Science in Computer Engineering, University of Guilan, 2020
- Master of Computer Science in Computer Science, University of New Brunswick, 2024

Publications: None

Conference Presentations: None