The Influence of Interdependence Relationships on Trust Repair Strategies and Collaboration Fluency

Alexandra Marcu[1]

Supervisor(s): Myrthe Tielman[1]**, Ruben Verhagen**[1]

[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 25, 2023

Name of the student: Alexandra Marcu
Final project course: CSE3000 Research Project
Thesis committee: Myrthe Tielman, Ruben Verhagen, Ujwal Gadiraju

Abstract

Intelligent agents are increasingly required to engage in collaboration with humans in the context of human-agent teams (HATs) to achieve shared goals. Interdependence is a fundamental concept in teamwork. It enables humans and robots to leverage their capabilities and collaboratively work towards a shared goal, fostering the development of trust through joint activities. Considering the great importance of trust, the effectiveness of trust repair strategies is crucial as they help mitigate the negative consequences of errors, enabling efficient collaboration between humans and robots. For this reason, the effectiveness of the trust repair strategies must be examined comprehensively by taking into consideration multiple factors, including the interdependence relationships within HATs. This paper aims to examine the impact of a mix of interdependence and independence relationships on trust violation and repair, but also on collaboration fluency. Thus, an experiment (n = 30) was conducted to study how interdependence affects trust violation, trust repair and collaboration fluency. Participants collaborated with a robot during a search and rescue mission in a simulated environment. Results show that there is a significant influence of interdependence on trust violations, but not on collaboration fluency or trust repair. Furthermore, the paper also emphasises the need for future research that investigates the effectiveness of trust repair strategies for HATs in different interdependence relationships.

1 Introduction

Human-agent teams (HATs) leverage the specific capabilities of each team member, resulting in improved performance across various domains [1–3]. In this collaborative approach, the team members not only perform individual actions but also engage in joint activities. The collaborative nature of HATs has generated significant interest in understanding the interaction between the components that enable and facilitate team success.

When it comes to teamwork, interdependence is a fundamental concept [4, 5], as it enables the team members to engage in joint activities and complement their abilities to overcome any individual limitations [6–8]. Several factors have been identified to be crucial when humans and robots are interdependent, such as trust, effective coordination and cooperation [9]. When team members lack specific capabilities, interdependence relationships established within the team can help resolve trust uncertainties and facilitate trust development [10].

Considering the great importance of trust in enabling effective collaboration in HATs, prior research has thoroughly analysed the factors that influence it [11, 12]. Recently, more studies have also started to explore how trust is affected by trust violations and ways to mitigate the negative effects of errors [13, 14]. Although prior studies have separately examined the effects of errors and interdependence on trust, there is little knowledge of how these two factors interact and influence trust dynamics. This knowledge gap has also been acknowledged in the context of trust repair in human teams, where the importance of analyzing contextual factors has been highlighted [15]. Although the current study focuses on human-agent teams, insights from related research emphasize the importance of examining the relationship between interdependence, trust violation and repair.

Given that there are multiple levels of interdependence, this paper will focus on investigating a mix of interdependence and independence relationships by answering the following research question:

> *How does a mix of interdependence and independence relationships influence the effectiveness of trust repair strategies?*

A mix of interdependence and independence relationships was chosen because it reflects a more realistic setting that could be encountered in HATs. While full interdependence scenarios provide valuable insights into trust dynamics, they may not capture the complexities found in real teams. To help answer this main research question, two secondary questions have been formulated:

1. *How does a mix of interdependence and independence relationships affect the trust violation?*
2. *How does a mix of interdependence and independence relationships affect the trust repair?*

Answering these questions could provide insights into the factors that influence trust dynamics within HATs. Currently, the understanding of trust repair strategies is limited to the type of error [14] or the level of risk [16]. However, this research could offer a more comprehensive and tailored approach to addressing trust violations. This could provide explanations on how these strategies work and when they are the most effective, allowing for adaptation to specific contexts.

On the other hand, effective coordination and cooperation are also important when humans and robots are interdependent. Fluency describes the coordination and the ease of collaboration with which team members perform joint activities [17]. As the demand for robots to engage in complex, interdependent joint activities with humans [6] continues to rise, the importance of collaboration fluency has also increased [18]. Currently, prior research does not agree on what factors affect collaboration fluency in HATs but states that it is particularly important in joint actions. For this reason, the secondary research question was found to be:

> *How does a mix of interdependent and independent relationships affect collaboration fluency in human-agent teams?*

The findings of this research have the potential to significantly enhance the understanding of the importance of developing autonomous agents that facilitate support for interdependence. Previous observations have highlighted the need for robots to not only improve their independent task performance, but also their capability to support interdependence in joint activities [6–8]. This is important as it encourages humans to perceive robots as teammates rather than tools.

This research paper is structured as follows. Section 2 introduces the background of the study and further explains concepts such as trust, interdependence relationships and collaboration fluency. Section 3 describes the experiment used to help answer the research questions. During this study, human participants collaborated with an agent during a search and rescue task in a simulated environment. This section includes a detailed overview of the design used, participant demographics, the task and the procedure. Section 4 gives some insights into the responsible research that has been applied during the experiments and the overall project. Section 5 reports the results obtained after conducting the experiment and section 6 interprets these results and identifies possible limitations of the user study. Finally, section 7 summarizes the most important findings.

2 Background

2.1 Trust and Trust Repair

Trust is a multidimensional concept and a fundamental factor that influences the success and effectiveness of collaborative activities, including those involving human-agent teams. Previous research has made significant efforts to understand and define trust complexities in the context of human-agent teams. A literature-based trust model identified primary elements as performance and moral trust, with dimensions including "reliable", "capable", "sincere" and "ethical" [19]. Additional studies propose dimensions such as ability, benevolence, integrity [20, 21], or predictability, reliability, persistence [22]. These variations highlight the complex nature of trust and the absence of consensus, contributing to the challenge of measuring and comprehending trust dynamics.

While it is important to understand the concept of trust, it is equally important to recognize the factors that influence it. Multiple aspects affect the trust of humans in robots, but the robot's performance characteristics such as reliability and competence appear to have the most significant contribution [11, 12]. The study [11] also recognises the moderate effect of environmental factors and also the lack of enough experiments that analyse the relationship with trust. As the interdependence level is related to environmental factors, there is a clear need for further research that analyse its influence on trust.

Trust has been demonstrated to play an essential role in enhancing task efficiency, accuracy and reliability, particularly in situations characterized by risk and uncertainty [10, 23]. Interactions involving risks raise concerns about the impact of errors on collaboration [13, 14]. To mitigate the consequences of trust violations caused by these errors, the agents can deploy trust repair strategies. However, prior research has mainly focused on the influence of the type of violation on different repair strategies, neglecting other factors or mechanisms [14]. Thus, it is necessary to address the gap in existing research by examining additional factors that can provide a comprehensive understanding of which repair strategies are most effective and in what conditions they should be applied.

2.2 Interdependence and Independence

Interdependence relationships lie at the core of teamwork and collaboration. Research suggests that in order to define the relationship between humans and robots as a "team", the element of interdependence needs to exist [5, 6]. In real-life situations, human-agent collaboration is most likely characterized by a mix of interdependence and independence relationships. Diverse forms of interdependencies arise when specific tasks require capacities that one team member may lack or recognise opportunities to be more efficient [24]. For example, in an urban search and rescue mission [25], required (hard) dependencies occur when an exploratory robot lacks the capacity to transport the victims it discovers. On the other hand, opportunistic (soft) dependencies arise when both humans and robots can remove an obstacle, but working together accelerates the process.

Recognizing the significance of these relationships in facilitating effective collaboration, researchers explored their impact on factors such as trust [10, 26]. They have identified that the level of interdependence within a team influences the extent of trust among team members. This is because interdependence provides opportunities to establish, develop, and maintain trust through repeated interactions and feedback [10]. However, these studies have there remains a gap in understanding how the influence of interdependence on trust within HATs is affected when errors occur. Given that these relationships are considered to be the mechanisms that support trust calibration [10], studying their influence on the repair strategies can help fill in this knowledge gap.

The need for future studies was also acknowledged by a study which aims to explain the impact of dependence relationships on trust repair for human teams [15]. While no prior work has specifically explored how interdependence relationships affect trust violation and repair in HATs, [25] could provide some foundation for this analysis. The study concludes that the level of interdependence is crucial in determining how different communication styles affect trust, so there is a need to include and explore the role of interdependence on other factors as well.

Previous studies have primarily focused on understanding how interdependence affects trust [7, 8], but little is currently known about how trust is developed and repaired over time, especially following trust violations. By investigating the impact of different trust repair strategies and considering the role of interdependence as a potential influencing factor, this research aims to provide insights into the mechanisms underlying trust repair in HATs.

2.3 Collaboration Fluency

In the context of interdependent relationships, collaboration fluency is particularly important as it refers to how well the teammates are synchronized and coordinated in performing the joint activity [17]. This is because robots are deployed more frequently in collaborative activities, where they are expected to coordinate fluently with humans. Research suggests that fluency in HATs stems from humans interacting with robots as they do with other people rather than with tools [27]. Interdependent relationships are considered to help robots be seen as teammates rather than tools, as people perceive themselves to be more in a team relationship [5]. This perspective indicates that fostering these relationships can have a substantial impact on fluency.

Conversely, human-agent collaboration is more difficult to describe and explain in the case of independence, when there is an absence of joint tasks. Continuous fluent interactions are considered to be necessary when engaging in joint actions, but there is little knowledge of the importance of fluency when the teammates perform some or all tasks independently. Additionally, existing research has not explored the difference in fluency between human-agent teams engaged in interdependent relationships, independent relationships or a combination of both. Given that many HATs operate in situations that fall between complete interdependence or independence, it becomes essential to examine how fluency is affected in these mixed scenarios. This aims to bridge the knowledge gap regarding collaboration fluency when the team is not characterized only by interdependent relationships.

3 Methods

3.1 Design

To understand how the mix of independent and interdependent relationships affects the trust repair strategy and collaboration fluency, we conducted an experiment. It followed a 3 x 5 mixed design with the interdependence level as the between-subject independent variable and time as a within-subject independent variable. The mixed interdependence condition was compared to a baseline condition representing full independence, where no interdependence relationships existed. For both conditions, the same trust repair strategy was used which involved expressing regret and explaining why the violation occurred. This was discovered to be the most effective for competence-based violations [28].

3.2 Participants

We recruited 30 participants from personal contacts (13 females and 17 males), 15 for each one of the conditions. One participant had an age range of 25-34 years old and the rest of the 29 participants were between 18-24 years old. In terms of education, 24 participants were high school graduates, while 6 participants had obtained a Bachelor's degree. Regarding gaming experience, three participants had no prior experience, nine participants had a little, six participants had a moderate amount, four participants had a considerable amount, and eight participants had a lot. Each participant signed an informed consent form before participating in the study.

3.3 Hardware and Software

The experiments were run on a laptop, which was used to launch and access the two-dimensional simulated search and rescue task. The environment was built using the Human-Agent Teaming Rapid Experimentation[1] software (MATRX), and it is similar to the one used in a previous study [25].

[1]MATRX Software: https://matrx-software.com/

3.4 Environment

The experiment was conducted using a search and rescue environment, implemented in MATRX. As shown in Figure 1, the world consisted of 14 distinct areas, containing elements such as 26 collectable objects, including critically injured victims (red), mildly injured victims (yellow) and healthy victims (green). Additionally, there were 12 obstacles, such as stones, rocks, and trees, as well as a drop zone. To simulate extreme weather conditions, flooded water was incorporated into the environment, reducing the agents' speed as they moved through it. The mission's goal is to locate and rescue all 8 critically and mildly injured victims within a 10-minute timeframe, after which the task is terminated and the objective metrics are saved. The search and rescue task was chosen because it allows for the manipulation of interdependence conditions, making it appropriate for studying the effects of different levels of interdependence on trust and collaboration.

Figure 1: The map represents the "God" view of the MATRX world used for this study. The right area of the world shows the drop zone with eight victims to search and rescue. This figure shows the starting position of the RescueBot and the human avatar.

3.5 Task

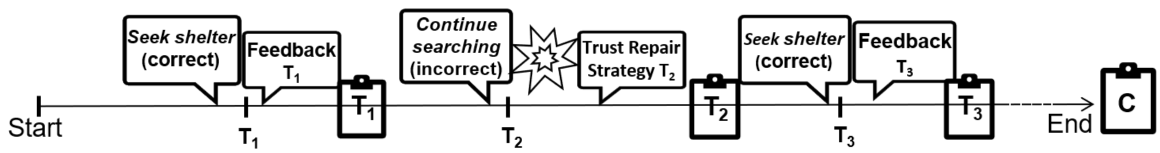
The main focus of the study is on how a mix of interdependence and independence relationships influence the trust repair strategy, compared to a baseline condition in which the human and agent are fully independent. This is done by studying the dynamic of trust before and after a trust violation and comparing how the decrease and increase in trust are affected by the interdependence condition. Therefore, it is essential to define these conditions and examine how they shape the interactions between humans and agents.

In the baseline condition, the human and robot worked independently without collaborating on any tasks. The mixed interdependence condition involved a combination of tasks

Table 1: Overview of messages sent by the RescueBot throughout the experiment. Parts of the text have been highlighted to enhance participants' comprehension of the messages.

Message Type	Message Content
Advice T_1	I have detected **extreme rain** arriving soon and predict it will cause new floods, so I advise you to **take shelter** in one of the areas as soon as possible and until the rain is over.
Feedback T_1	My advice was **correct**, that weather was extreme! If you had not taken shelter, you would have lost important mission time due to injuries and 10 points of our score.
Advice T_2	I have detected **light rain** arriving soon but predict it will cause no floods, so I advise you to **continue searching and rescuing** victims.
Repair T_2	My advice was **wrong**. The amount of rain was heavy instead of light and because of that, my flood prediction was incorrect. **I am really sorry**.
Advice T_3	I have detected **extreme rain** arriving soon and predict it will cause new floods, so I again advise you to **take shelter** in one of the areas as soon as possible and until the rain is over.
Feedback T_3	My advice was **correct** now, that weather was extreme! If you had not taken shelter, you would have lost important mission time due to injuries and 10 points of our score.

Figure 2: Schematic timeline of the experiment. Each phase consisted of advice from the agent, a feedback message or a feedback message using the trust repair strategy and a trust questionnaire (clipboard icon with the letter T). The questionnaire at the end of the experiment measured collaboration fluency (clipboard icon with the letter C).



that have different levels of interdependence. The condition included tasks without constraints which could be performed by either team member alone, such as carrying mildly injured victims. Additionally, removing rocks was only possible if the human and robot worked together, establishing a required interdependence between them. On the other hand, the removal of stones could be done either individually or collaboratively, resulting in an opportunistic interdependence as performing the action together was more efficient. Finally, certain tasks required independence, where specific roles were assigned to either the human (e.g. carrying critically injured victims) or the robot (e.g. removing trees). This combination of tasks with different interdependence levels formed the basis of the mixed interdependence condition. For the remainder of this paper, we will refer to the scenario that combines interdependence and independence relationships as the "mixed interdependence condition".Conversely, the term "baseline condition" will be used to describe the scenario where there is full independence between the team members.

3.6 Procedure

The participants were first instructed to read and fill out the informed consent form. After the participants were assigned to one of the interdependence conditions, they followed a tutorial to get familiar with the environment, controls and messaging system. The interdependence condition used during the tutorial was the same for all participants, regardless of their assignment for the official task.

For the official task, the participants were instructed to collaborate with the RescueBot during a search and rescue mission in a town that suffered from extreme weather. Additionally, they were made aware of the possibility of further bad weather hitting the town, which could potentially cause harm to the human participant and result in penalties. They were also informed that they should pay attention to the chat as the RescueBot will send warnings about the weather. A notification sound was played every time a message which required immediate attention from the participant was sent. During the mission, the participants received warnings indicating whether they should take shelter or continue with the task, based on the predicted weather.

The schematic timeline for the experiment that illustrates the sequence of events is depicted in Figure 2. There were a total of 3 warning messages, sent at an interval of 2 minutes from each other. The first and the third messages were correct and if participants followed the advice and sought shelter, no penalties were applied. The second message was incorrect, resulting in a time and score penalty for the player. The purpose of applying the penalty was to cause a trust violation. To repair trust, the agent implemented a repair strategy of expressing regret and providing a detailed explanation for why the trust violation occurred. An overview of all messages sent by the robot can be found in Table 1. After each feedback message, the game paused, and participants were instructed to switch to a different tab to complete a questionnaire about their perceived trust in the RescueBot. Once the participants finished the mission or the 10 minutes elapsed, they were asked to fill out an additional questionnaire regarding the collaboration fluency they experienced during the mission.

3.7 Measures

To analyze the influence of the mix of independence and interdependence relationships on the trust repair strategy and on collaboration fluency, a set of objective and subjective measures was used. All objective data was automatically logged using MATRX and the questionnaire responses were collected using Qualtrics.

3.7.1 Subjective Measures

We subjectively measure trust and collaboration fluency using questionnaires. Trust was measured using a 5-point Likert scale proposed by [29], with a total of 8 items. The questionnaire included questions about predictability (i.e. *"The outputs of RescueBot are very predictable"*), reliability (i.e. *"RescueBot is very reliable. I can count on it to be correct all the time."*) and efficiency (i.e. *"RescueBot is efficient in that it works very quickly."*).

The second subjective measure is collaboration fluency and it was evaluated using an 8 items scale developed by [17], which assesses different facets of collaborative performance using a 7-point Likert scale. Questions about commitment to success (i.e. "*The robot was committed to the success of the team*") or fluency (i.e. "*The human-robot team worked fluently together*", "*The robot contributed to the fluency of the collaboration*") were included. For both measures, the scores were converted to a final numeric value by calculating the mean of the responses.

3.7.2 Objective Measures

Even though trust and collaboration fluency are often measured subjectively, they can also be inferred from human behaviour. To help analyse the results for trust, we measured advice acceptance and the percentage of joint actions. Advice acceptance recorded whether the human was located in a shelter during the bad weather. A single questionnaire item was also used to address accidental shelter usage. According to prior work, collaborative actions help develop and maintain trust [10]. Thus, the percentage of joint actions before and after the trust violation was recorded as it could motivate higher trust.

While there is no accepted set of metrics for fluency in HATs, the use of objective measures for fluency was informed by prior research that indicated their relevance. For this research, only robot idle time was objectively measured, as it has been previously used as a metric to evaluate team fluency [17]. Performance metrics such as completeness, score and total task time were also recorded as they might help explain the subjective results. While fluency in HATs has been associated with improved team performance [30, 31] and shorter total task time [18, 32], these performance metrics have not been used as measures of collaboration fluency. Completeness was calculated by dividing the number of rescued victims by the total number to be saved. The score was computed by adding 3 points for all the mildly injured victims and 6 points for all the critically injured victims that have been rescued within the mission time. The total time represented duration (in ticks) from the start of the game until the mission is completed or the 10-minute time limit expires. Finally, idle time records the percentage of time the RescueBot did not move its position.

4 Responsible Research

The reproducibility factor is an important aspect that must be taken into consideration and reflected upon for responsible research. Moreover, as an user experiment was conducted to answer the research questions, it is also important to look into the ethical concerns.

To ensure reproducibility, the entire codebase utilized in the experiments is accessible via a public fork of the research institution's sub-repository [2]. Given that multiple types of objective measures were recorded for both mixed interdependence and baseline conditions, it was essential to maintain consistency in order to facilitate accurate comparisons between the different conditions. While thorough reviews were conducted on the code related to logging the measurements, it was discovered that certain updates to the codebase resulted in an improper recording of one of the logs for half of the participants. In compliance with responsible research practices, the affected measurement was removed from the logging file, thus from the results and the analysis. Its inclusion would not have yielded accurate or meaningful outcomes. Moreover, this research ensures reproducibility by including automated data analysis utilizing the R programming language, the publication of all gathered data such that anyone could inspect it and a detailed explanation of the experimental setup. It is important to also note that the data from all 30 participants was included in the analysis.

The research study obtained approval from the Human Research Ethics Committee (HREC) at TU Delft, ensuring compliance with ethical considerations due to its minimal-risk nature. To collect subjective measurements and personal data, the survey tool *Qualtrics* [3] was employed, known for its compliance with privacy laws (GDPR). While the study is not anticipated to pose known risks, it is important to acknowledge the possibility of a data breach in the online environment. To prioritize privacy and confidentiality, the collected data is anonymized and limited to non-sensitive information (gender, age range, education level and gaming experience). These measures also minimize the risk of re-identification. Another potential risk of the experiments is bias due to participants being recruited from the researchers' personal network, potentially creating pressure to behave or respond according to the researchers' expectations. To address this, participants received informed consent following TU Delft's guidelines and it was emphasized the importance of honest responses to ensure unbiased data collection.

5 Results

5.1 Trust

For the dependent variable Trust, a Mixed ANOVA test was used with the between-factor Interdependence (baseline and mixed interdependence) and the within-factor Time (prior to violation [T_1], versus after violation [T_2] versus after repair [T_3]). As the data met all assumptions, the Mixed ANOVA test was used in the analysis.

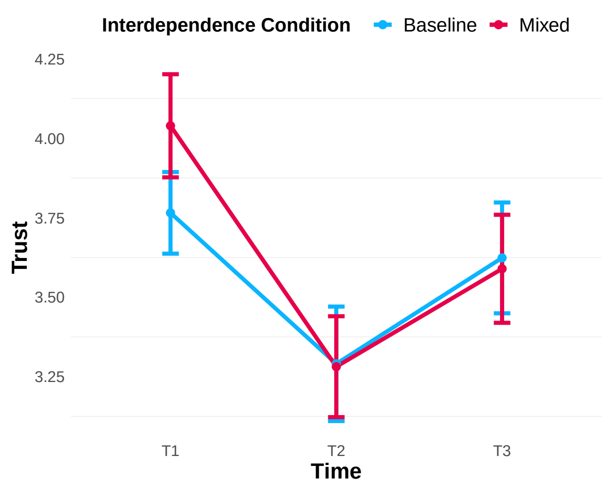
---

[2] GitHub repository used for this project: https://github.com/mawakeb/CSE3000-2023-trust-repair

[3] Qualtrics software: https://www.qualtrics.com/

A significant main effect of Time on trust was found ($F(2, 84) = 7.14$, $p = 0.001$). Considering the Bonferroni adjusted p-value, the simple main effect of time was significant for the mixed interdependence ($p < 0.05$). Although the effect was not significant for the baseline condition ($p = 0.062$), it approached significance. Results of the post-hoc test show that the mean trust value was significantly different between [T_1 - T_2] for mixed interdependence ($p < 0.05$), but not for the baseline ($p = 0.11$) as well. After the trust repair strategy was used [T_2 - T_3], no significant differences were observed in the mean trust for either interdependence condition ($p = 0.09$ for mixed interdependence and $p = 0.076$ for baseline). Finally, there were also no significant differences between the initial and the final trust values [T_1 - T_3] ($p = 0.056$ for mixed interdependence and $p = 1$ for baseline). These results can be viewed graphically in Figure 3.

Figure 3: Trust values over time for [T_1] (before violation), [T_2] (after violation), [T_3] (after repair strategy). The values are displayed for the baseline and mixed interdependence levels.



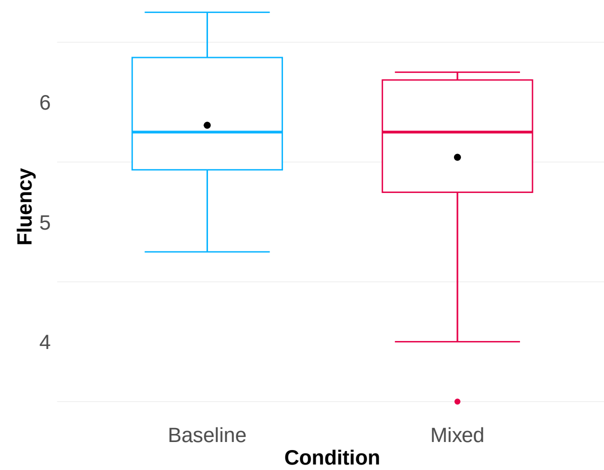
5.2 Joint Actions

For the mixed interdependence, the percentage of joint actions before and after the trust violation was analysed using a paired t-test. Before the trust violation, the number of joint actions ($M = 0.411$, $SD = 0.28$) was not found significantly higher than after the violation ($M = 0.292$, $SD = 0.147$), $t(14) = -1.35$, $p = 0.19$.

5.3 Collaboration Fluency

Given that the normality assumption was not met for the mixed interdependence condition as assessed by Shapiro-Wilk's test of normality ($p < 0.05$), a Mann-Whitney test was used. The median fluency value for both the baseline and mixed interdependence conditions was 5.75, with an interquartile range (IQR) of 0.938. The Mann-Whitney test showed no significant differences ($W = 132$, $p = 0.42$) between the two conditions (see Figure **??**).

Figure 4: Box-plots for comparing fluency between the mix of independence and interdependence relationships (Mixed) and the baseline.



5.4 Performance

To analyse the total task time between the two conditions, a Mann-Whitney test was used. This was because the normality assumption was not met for the mixed interdependence condition, as assessed by Shapiro-Wilk's test of normality ($p < 0.05$). The total task time for the mixed interdependence condition was found significantly greater than for the baseline ($W = 49$, $p = 0.008$).

The Mann-Whitney test was used to analyze both score and completeness as the normality assumption was not met for both variables. The results show no significant difference in terms of score ($W = 131.5$, $p = 0.43$) or completeness ($W = 143$, $p = 0.11$) between the 2 interdependence conditions.

5.5 Ratio of Idle Time

The ratio of total idle time of the robot was analysed using the Independent-samples t-test. Despite Shapiro-Wilk's test of normality indicating that the baseline condition does not follow a normal distribution ($p = 0.046$) it was decided to proceed with the parametric test. This was based on the slight deviation from the threshold ($p = 0.05$). The total idle time for the baseline condition ($M = 0.352$, $SD = 0.084$) was significantly lower compared to the mixed interdependence condition ($M = 0.535$, $SD = 0.103$), $t(28) = -5.3$, $p < 0.001$.

6 Discussion

6.1 Trust

In the **trust violation** phase [T_1-T_2], the results indicate a significant decline in trust for the mixed interdependence condition, but not for the baseline. A prior study [25] suggests that higher interdependence leads to a more critical evaluation of trust due to the perceived impact of errors on performance. Thus, the higher the interdependence level, the more drastically the trust is expected to decrease. While research argues that interdependence should have a more positive impact on how trust is maintained [15, 33, 34], the argument from [25] can help explain the difference in results in the context of errors. Another interpretation could be that trust can only be damaged when a person has developed some higher level of trust, prior to the error [15]. Although there is not a significant difference in the initial trust between the two conditions, the trust values for the mixed interdependence condition seem to be higher than for the baseline. In the context of trust violations in human-human teams, [15] argues that the higher the interdependence, the less significant the trust decrease. This is due to human's inclination to dismiss negative outcomes from people that they trust as they do not see the errors as a trust violation. The difference in results could be attributed to the fact that the participants did not perceive working with the robot the same as working with another human. Although prior work claims that a high degree of interdependence can help the robots be viewed as teammates rather than tools [5], our results seem to contradict this in the case of trust violations.

The results for **trust repair** [T_2-T_3] do not align with prior work that highlights the importance of interdependence relationships in developing and maintaining trust [10]. While it was expected that the lack of joint actions in the baseline condition would not significantly increase trust, the required and opportunistic dependencies from the mixed interdependence condition were anticipated to facilitate trust repair as people collaborated more closely with the robots. One reason for the limited impact of the trust repair strategy in the mixed condition could be attributed to the collaboration not being exclusively based on joint actions. Additionally, it is important to note that previous research did not account for cases involving trust violations, which could contribute to the observed differences in results.

Additionally, the results for the trust repair also contradict a prior study suggesting that apologies including expressions of regret were most effective in repairing trust after a violation [28]. As the experiment did not involve any joint actions, the mission can be considered more similar to the baseline. Another critical point is that in that study, the repair strategy was deployed after the second trust value [T_2] was recorded, while in the current experiment, it was recorded just after. This could mean that the second value of trust from our results are somewhat higher than the ones from [28], as it was already manipulated by the trust repair strategy.

Analyzing trust dynamics in relation to the percentage of joint actions can offer new perspectives for the obtained results. According to prior research [34], trust is developed through repeated interactions and feedback between humans and robots. In the mixed interdependence condition, a larger portion of actions was performed individually during the game due to the limited number of required and opportunistic dependencies. Additionally, the limited number of joint actions may have resulted in a higher perceived interdependence between [T_1-T_2], but a lower one between [T_2-T_3]. This may explain the similar outcomes between the interdependence conditions after the trust violation and the differences prior to it.

6.2 Collaboration Fluency

For collaboration fluency, the results of the subjective measurements show that different interdependence conditions do not influence it significantly. While the subjective measures indicated no influence of interdependence on fluency, the objective measure provided a different perspective. The results show that the baseline had a significantly lower robot idle time than the mixed interdependence condition. Previous research [17] indicates that this could mean a higher degree of collaboration fluency. However, it is important to recognize that the robot's idle time is influenced by various factors such as the participant's level of attention, the number of messages exchanged, and the number of joint actions. Therefore, it is more challenging to directly relate this metric to fluency in the current experiment. The same reasoning applies to the significant difference in total task time.

Fluency is considered to be associated with better team performance [18,35]. As neither score nor completeness showed significant differences between the interdependence conditions, we can not conclude that they could have influenced subjective fluency.

Prior research mentions that increasing interdependence also increases the need for better coordination and collaboration [24], which could result in a more difficult mission [25]. This is in line with the results, as the complexity of the mission task for the mixed interdependence condition, which required participants to frequently adapt how they collaborated with the robot based on the specific task, lowered the subjective fluency values.

6.3 Limitations

The presented study has several limitations. First, the participants were all students recruited from TU Delft, which means that they are more familiar with technology than an average user who would have contact with robots. Moreover, the homogeneity of the group of participants might have influenced how generalizable the results are.

Secondly, the visibility of the advice, feedback and repair messages could have also influenced the results. Even though the experiment was designed to help participants acknowledge the messages as fast as possible, there was always the risk that some participants did not read the message. This resulted in trust violations even when it was not caused by the robot due to the fact that the participants were not able to seek shelter in time. Moreover, since there are no means of verifying whether participants actually read the repair message, it is possible that the trust repair was solely influenced by the interdependence relationships rather than the strategy used.

Finally, the second recorded value of trust was recorded after the message aiming to repair the trust was sent, allowing the participant to read it before filling in the questionnaire. For this reason, the trust values at [T_2] might be already influenced by the trust repair strategy, making the results more difficult to compare to other studies.

7 Conclusion and Future Work

This research project aimed to investigate the impact of a mix of interdependence and independence relationships on trust violation, trust repair and collaboration fluency within HATs. By addressing the first research question, this study highlights the importance of further examining the influence of interdependence relationships on the effectiveness of trust repair strategies. The findings indicate a significant impact of a mix of interdependence and independence relationships on trust violation, resulting in lower trust values. The negative effect of these relationships emphasises that their effects are influenced by the occurrence of errors during the collaboration. Moreover, the results indicate a slow recovery of trust after a trust violation in both interdependence conditions, suggesting a negative influence on the effectiveness of the trust repair strategy. Future studies on trust repair strategies that analyse the level of interdependence as a mediating factor are crucial as effective utilization of these strategies can determine the collaboration efficiency in HATs. All these results highlight the difference in how humans establish and develop trust in the case of competence-based errors compared to the cases when no violations occur. Moreover, they also give insights into how people perceive and assess trust violations depending on how much they need to directly collaborate and rely on the agent.

The second research question explored the influence of a mix of interdependent and independent relationships on collaboration fluency. The results reveal that collaboration fluency is not significantly influenced by the level of interdependence. On the other hand, the robot idle time was significantly affected, but a more careful consideration revealed that the nature of the collaborative task could have had a much greater influence.

Overall, the results show that there are important differences between distinct levels of interdependence on trust violation, but not on trust repair or collaboration fluency. Interdependence has the potential to shape the collaboration dynamics within HATs, which emphasises the need for further studies into more specific cases of interdependence, such as fully required or opportunistic relationships. By exploring these variations, future research can provide a more complete understanding of the influence of interdependence on trust violation, repair, and collaboration fluency within HATs.

References

[1] C. E. Bartlett and N. J. Cooke, "Human-Robot Teaming in Urban Search and Rescue," *Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting*, vol. 59, pp. 250–254, 9 2015.

[2] C. Culbert, J. Rochlis, F. Rehnmark, D. Kortenkamp, C. Weisbin, and G. Rodriguez, "Activities of the NASA Exploration Team Human-Robotics Working Group," 9 2003.

[3] N. Robinson, J. Williams, D. Howard, B. Tidd, F. Talbot, B. Wood, A. Pitt, N. Kottege, and D. Kulić, "Human-Robot Team Performance Compared to Full Robot Autonomy in 16 Real-World Search and Rescue Missions: Adaptation of the DARPA Subterranean Challenge," *arXiv (Cornell University)*, 12 2022.

[4] E. Salas, N. J. Cooke, and M. J. Rosen, "On Teams, Teamwork, and Team Performance: Discoveries and Developments," *Human Factors*, vol. 50, pp. 540–547, 6 2008.

[5] C. Nass, B. J. Fogg, and Y. Moon, "Can computers be teammates?," *International Journal of Human-Computer Studies*, vol. 45, pp. 669–678, 12 1996.

[6] "Autonomy and interdependence in human-agent-robot teams," *IEEE Intelligent Systems*, vol. 27, no. 2, pp. 43–51, 2012.

[7] J. D. Bradshaw, V. Dignum, C. M. Jonker, and M. Sierhuis, "Human-agent-robot teamwork," *IEEE Intelligent Systems*, vol. 27, pp. 8–13, 3 2012.

[8] M. P. Johnson, J. M. Bradshaw, P. J. Feltovich, C. M. Jonker, B. Van Riemsdijk, and M. Sierhuis, *The Fundamental Principle of Coactive Design: Interdependence Must Shape Autonomy*. 5 2010.

[9] E. Salas, D. E. Sims, and C. S. Burke, "Is there a "Big Five" in Teamwork?," *Small Group Research*, vol. 36, pp. 555–599, 10 2005.

[10] M. Johnson and J. M. Bradshaw, "Chapter 16 - the role of interdependence in trust," in *Trust in Human-Robot Interaction* (C. S. Nam and J. B. Lyons, eds.), pp. 379–403, Academic Press, 2021.

[11] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. de Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 53, no. 5, p. 517–527, 2011.

[12] B. M. MUIR and N. MORAY, "Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation," *Ergonomics*, vol. 39, no. 3, pp. 429–460, 1996. PMID: 8849495.

[13] M. Salem, G. Lakatos, F. Amirabdollahian, and K. Dautenhahn, "Would You Trust a (Faulty) Robot?," 3 2015.

[14] C. Esterwood and L. P. Robert, "A literature review of trust repair in hri," in *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 1641–1646, 2022.

[15] E. C. Tomlinson, "The context of trust repair efforts: Exploring the role of relationship dependence and outcome severity," *Journal of trust research*, vol. 1, pp. 139–157, 10 2011.

[16] P. Robinette, A. M. Howard, and A. R. Wagner, "Effect of Robot Performance on Human–Robot Trust in Time-Critical Situations," *IEEE Transactions on Human-Machine Systems*, vol. 47, pp. 425–436, 8 2017.

[17] G. Hoffman, "Evaluating Fluency in Human–Robot Collaboration," *IEEE Transactions on Human-Machine Systems*, vol. 49, pp. 209–218, 4 2019.

[18] C. Chao and A. L. Thomaz, "Timing in Multimodal Turn-Taking Interactions: Control and Analysis Using Timed Petri Nets," *Journal of human-robot interaction*, pp. 4–25, 8 2012.

[19] B. F. Malle and D. Ullman, "Chapter 1 - a multidimensional conception and measure of human-robot trust," in *Trust in Human-Robot Interaction* (C. S. Nam and J. B. Lyons, eds.), pp. 3–25, Academic Press, 2021.

[20] C. Esterwood and L. P. Robert, "Do you still trust me? human-robot trust repair strategies," *2021 30th IEEE International Conference on Robot amp; Human Interactive Communication (RO-MAN)*, 2021.

[21] C. Esterwood and L. P. R. Jr, "Three strikes and you are out!: The impacts of multiple human-robot trust violations and repairs on robot trustworthiness," *Computers in Human Behavior*, vol. 142, p. 107658, 2023.

[22] A. Uggirala, A. K. Gramopadhye, B. J. Melloy, and J. E. Toler, "Measurement of trust in complex and dynamic systems using a quantitative approach," *International Journal of Industrial Ergonomics*, vol. 34, no. 3, pp. 175–186, 2004.

[23] R. Mayer, J. Davis, and F. D. Schoorman, "An Integrative Model Of Organizational Trust," *Academy of Management Review*, vol. 20, p. 730, 7 1995.

[24] M. Johnson, J. M. Bradshaw, P. J. Feltovich, C. M. Jonker, M. B. Van Riemsdijk, and M. Sierhuis, "Coactive design: Designing support for interdependence in joint activity," *Journal of Human-Robot Interaction*, vol. 3, no. 1, p. 14–15, 2014.

[25] R. S. Verhagen, M. A. Neerincx, and M. L. Tielman, "The influence of interdependence and a transparent or explainable communication style on human-robot teamwork," *Frontiers in Robotics and AI*, vol. 9, 2022.

[26] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. F. Camerer, "Not so different after all: A Cross-Discipline view of trust," *Academy of Management Review*, vol. 23, pp. 393–404, 7 1998.

[27] M. Paliga, "Human–cobot interaction fluency and cobot operators' job performance. The mediating role of work engagement: A survey," *Robotics and Autonomous Systems*, vol. 155, p. 104191, 9 2022.

[28] E. S. Kox, J. H. Kerstholt, T. F. Hueting, and P. W. de Vries, "Trust repair in human-agent teams: The effectiveness of explanations and expressing regret," *Autonomous Agents and Multi-Agent Systems*, vol. 35, no. 2, 2021.

[29] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman, "Measures for explainable AI: Explanation goodness, user satisfaction, mental models, curiosity, trust, and human-AI performance," *Frontiers in computer science*, vol. 5, 2 2023.

[30] G. Hoffman and C. Breazeal, "Effects of anticipatory perceptual simulation on practiced human-robot tasks," *Autonomous Robots*, vol. 28, pp. 403–423, 12 2009.

[31] S. Nikolaidis, D. Hsu, and S. S. Srinivasa, "Human-robot mutual adaptation in collaborative tasks: Models and experiments," *The International Journal of Robotics Research*, vol. 36, pp. 618–634, 2 2017.

[32] A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa, "Effects of Robot Motion on Human-Robot Collaboration," 3 2015.

[33] A.-S. Ulfert and E. Georganta, "A Model of Team Trust in Human-Agent Teams," 10 2020.

[34] T. O'Neill, N. McNeese, A. Barron, and B. Schelble, "Human–autonomy teaming: A review and analysis of the empirical literature," *Human Factors*, vol. 64, no. 5, pp. 904–938, 2022. PMID: 33092417.

[35] A. N. Vazquez, *Evaluating Team Fluency in Human-Industrial Robot Collaborative Design Tasks*. 1 2022.