Designing Beyond Anthropomorphism: Human- and Robot-Like Error Recovery for Robot Navigation Errors

Hideki Garcia Goo
University of Twente
Enschede, Netherlands
h.garciagoo@utwente.nl

Bob Schadenberg
University of Twente
Enschede, Netherlands
bob.schadenberg@nipv.nl

Sebastian Schneider
University of Twente
Enschede, Netherlands
s.schneider@utwente.nl

Vanessa Evers
Centrum Wiskunde & Informatica
Amsterdam, Netherlands
University of Twente
Enschede, Netherlands
Nanyang Technological University
Singapore, Singapore
Vanessa.Evers@cwi.nl

Abstract

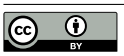
Robots operating in shared human environments will inevitably make mistakes. These errors can negatively affect users' attitudes toward robots, making effective recovery strategies essential. While prior work has focused on human-like repair behaviors, these may be mismatched with a robot's capabilities and increase perceptions of its social intelligence. It remains unclear whether robot-like recovery strategies can restore trust without increasing perceived social intelligence. We conducted an online animated video study ($N = 219$) in which participants observed a healthcare robot making a navigation error followed by one of four recovery strategies: none, robot-like (functional), human-like (socio-emotional), or combined. Trust and perceived social intelligence (PSI) were measured before and after the error. Both trust and PSI decreased significantly following the error. Recovery strategies restored both measures, although their effects differed. Robot-like recovery restored trust more effectively than human-like recovery while avoiding increases in PSI, whereas human-like recovery elevated PSI beyond baseline levels. The combined strategy produced the strongest overall recovery. These findings suggest that robot-like recovery are an effective approach to trust repair while avoiding the increases in perceived social intelligence associated with human-like recovery behaviors.

CCS Concepts

• **Human-centered computing** → **Empirical studies in interaction design**; *Interaction techniques*; • **Computer systems organization** → *Robotics*.

Keywords

robot, error, trust, recovery strategies, mitigation strategies, healthcare, navigation



This work is licensed under a Creative Commons Attribution 4.0 International License.
IVA 2026, Puebla, Mexico
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2647-7/26/09
https://doi.org/10.1145/3806774.3827977

ACM Reference Format:
Hideki Garcia Goo, Bob Schadenberg, Sebastian Schneider, and Vanessa Evers. 2026. Designing Beyond Anthropomorphism: Human- and Robot-Like Error Recovery for Robot Navigation Errors. In *ACM International Conference on Intelligent Virtual Agents (IVA 2026), September 07–11, 2026, Puebla, Mexico.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3806774.3827977

1 Introduction

Advances in robotics have enabled robots to operate in shared human environments, where they must navigate complex and unpredictable situations where errors are inevitable. These errors require robots to communicate with users in a responsive manner [29, 44]. Prior research in Human–Robot Interaction (HRI) shows that robot errors influence users' attitudes, affecting trust and acceptance of the robot [24, 41]. These effects depend not only on the error itself but also on how users perceive it. Performance errors (caused by technical malfunctions) can reduce perceived competence, while social errors (arising from inappropriate social behavior) can negatively affect perceptions of a robot's socio-affective abilities [40]. However, this distinction is often unclear, as users may interpret performance failures as social mistakes and vice versa.

To mitigate the negative impact of errors, most HRI research adapts human repair strategies such as apologies, explanations, humor, or promises to robotic systems [5, 15, 23, 25, 37, 41, 46, 48, 49]. While human-inspired responses can support trust recovery [23, 31, 46], their effectiveness is inconsistent [7, 8, 13, 28]. Moreover, such strategies may inflate perceptions of a robot's social intelligence (especially for service robots whose role is not inherently social), shaping users' expectations of its capabilities [42]. Robot apologies have also been shown to influence how responsibility for errors is attributed, potentially shifting responsibility away from designers or operators [27]. Together, these limitations motivate the exploration of alternative robot-like recovery approaches.

Alternative approaches that leverage robot-specific communication, such as system-style error messages or signals through lights and sounds, remain comparatively underexplored in HRI [17,

22, 37, 38]. Beyond aligning with ethical guidelines that recommend avoiding unnecessary anthropomorphism [3, 32], robot-like recovery strategies may better communicate a robot's actual capabilities while avoiding inflated social expectations [42]. Despite these potential advantages, little is known about how robot-like and human-like recovery strategies differentially influence trust and perceived social intelligence. Understanding both outcomes is important because recovery behaviors may restore trust while simultaneously shaping how socially capable users perceive the robot to be.

To address this gap, we examine how navigation errors made by a healthcare delivery robot (see Figure 2) affect *trust* and *perceived social intelligence*, and whether different responses following an error can mitigate these effects. Using an online video-based study with animated interactions between a virtual robot and a virtual human, we compare a *human-like* response (a non-verbal expression of regret), a *robot-like* response (communicating the robot's ability to learn from the error), their combination, and a no-recovery condition. A short version of this work has previously been published as a late-breaking work focusing only on trust outcomes [18].

This work makes two key contributions. First, it examines both trust and perceived social intelligence, providing a broader understanding of users' responses to robot navigation errors. Second, it reveals how robot-like and human-like recovery strategies differentially influence trust and perceived social intelligence, informing the design of recovery behaviors for service robots.

2 Background

2.1 To err is robot

Prior work shows that robot errors generally have negative effects on user attitudes across different domains [12]. For example, service robots delivering incorrect food orders reduced participants' satisfaction [47], while errors in restaurant service robots negatively affected perceptions of politeness, competence, trust, and likability [28]. Similarly, navigation failures in home companion robots reduced perceived reliability and trustworthiness [36]. However, some studies report positive effects of robot errors [2, 23, 30]. These positive effects are typically observed in controlled or low-stakes settings, where imperfect behaviors can increase anthropomorphism and, in turn, perceived relatability and likability, consistent with the Pratfall Effect. [30, 35].

The impact of robot errors is not uniform. Severe failures typically produce stronger reactions and larger decreases in trust [34, 39]. Over longer interactions, repeated unmet expectations can create an "expectation gap," which has been identified as a key reason why people reject robots altogether [41]. Prior work also shows that the effects of robot errors depend on factors such as perceived anthropomorphism and violation type [8, 15].

Researchers have proposed taxonomies that distinguish between performance errors, where the robot fails to complete a task, and social errors, where it violates social norms [20, 21, 40, 41]. Performance errors tend to reduce perceptions of competence and trust, while social errors diminish perceived social intelligence [40]. However, these distinctions are not always clear from the user's perspective: technical failures can result in socially inappropriate behavior (e.g., invading personal space), leading users to interpret performance failures as social errors.

Taken together, these considerations suggest that common robot errors, such as navigation failures, may simultaneously affecting functional evaluations (e.g., trust) and social perceptions (e.g., perceived social intelligence).

2.2 To recover is design

Robot errors are commonly understood as violations of user trust, making trust recovery a central challenge in human–robot interaction. Trust develops when a robot behaves reliably and according to user expectations; when it makes a mistake or behaves unexpectedly, this trust can be violated [15].

To address these violations, prior work has explored a wide range of mitigation or recovery strategies aimed at restoring trust after robot failures. Common approaches include apologies, denials, explanations, and promises [15, 16, 24, 40, 41], although effectiveness varies depending on factors such as error severity, perceived anthropomorphism, and timing [4, 9, 10, 33], with no single strategy consistently outperforming others [4, 7, 9, 13, 23, 28, 33, 38]. Moreover, recovery strategies may not fully restore trust, as some strategies have not been shown to outperform a condition in which errors are avoided [11].

Most of these recovery strategies rely on human-like communication, raising both practical and ethical concerns. Although human-like strategies can effectively restore trust and anthropomorphic cues can benefit HRI [45, 47], previous research has shown the importance of aligning anthropomorphic cues with the robot's actual capabilities [42] to avoid misleading users on the robot's socio-affective abilities and functional capabilities. Also, the use of apologies as a recovery strategy has raised ethical concerns by potentially shifting perceived accountability away from developers or operators [27]. These limitations motivate the need to investigate recovery strategies that align more closely with robots' functional capabilities rather than relying solely on human-like responses.

3 Problem Statement

Despite growing research on robot error recovery, most work focuses on anthropomorphic responses such as apologies, expressions of regret, and promises. However, many service and assistive robots are not designed to simulate rich socio-affective behavior, and such strategies may inflate users' expectations or misrepresent the robot's capabilities. It therefore remains unclear whether recovery strategies that communicate a robot's functional capabilities (robot-like), rather than relying on anthropomorphic cues, can effectively restore trust while avoiding increases in perceived social intelligence. To address this gap, we examine how navigation errors affect trust and perceived social intelligence, and compare robot-like and human-like recovery strategies following such errors. We investigate whether robot-like recovery can restore trust without increasing perceived social intelligence, thereby providing guidance for designing recovery behaviors that align with a robot's intended capabilities.

RQ1: To what extent do navigation errors affect people's attitudes toward a healthcare delivery robot?

Prior work shows that robot errors reduce trust, competence, and reliability across service contexts [28, 36, 47]. Moreover, because navigation errors can be interpreted as both performance and social failures, they are expected to reduce both trust and perceived social intelligence [40]. **H1a:** Participants' trust and perceived social intelligence will decrease following a navigation error.

Previous work suggests that more severe robot failures elicit stronger negative user reactions and larger decreases in trust [34, 39]. **H1b:** Greater perceived error severity will be associated with larger decreases in trust and perceived social intelligence.

RQ2: To what extent can a healthcare delivery robot recover from navigation errors using different recovery strategies?

Prior work demonstrates that recovery strategies influence trust following robot failures, although no single strategy consistently outperforms others across contexts [16, 26, 41]. At the same time, anthropomorphic behaviors can improve users' perceptions of robots [45], but have also been argued to increase ethical risks when they imply capabilities the robot does not possess [42]. These findings suggest that robot-like and human-like recovery strategies may affect different aspects of users' evaluations on trust and perceived social intelligence. **H2a:** The type of recovery strategy will influence recovery, such that robot-like strategies will lead to greater recovery of trust, whereas human-like strategies will lead to higher perceived social intelligence.

Previous work has shown that more severe robot failures lead to larger decreases in trust [4, 20]. Failure severity has also been found to influence the effectiveness of recovery strategies, with users responding differently to recovery attempts following more versus less severe failures [4, 48]. Together, these findings suggest that larger initial declines in trust and perceived social intelligence may be more difficult to recover from. **H2b:** Larger initial decreases in trust and perceived social intelligence will be associated with less recovery following mitigation.

RQ3: To what extent is the effectiveness of mitigation strategies influenced by user and error characteristics?

User characteristics and interaction context can influence responses to robot behavior [6, 23]. **H3a:** Participant characteristics (e.g., gender, age, experience with robots, healthcare background) will influence the effectiveness of recovery strategies.

Previous research also shows that error severity can affect the effectiveness of mitigation strategies [4, 9, 48]. **H3b:** Manipulated error severity (i.e., error condition) will influence the effectiveness of recovery strategies.

These hypotheses are tested in an online video-based study in which participants observe navigation errors of varying severity and mitigation strategies, with repeated measures of trust and perceived social intelligence.

4 Method

4.1 Participants

Participants were recruited through Prolific, a crowd-sourcing platform, and were compensated €8 for their participation. The survey took an average of 20.38 minutes (*SD* = 12.91). Additional inclusion criteria required participants to be proficient in English and to have access to a computer with audio capabilities. This study was approved by the Ethics Committee Computer & Information Sciences of the University of Twente (reference number: #230698).

The study included 219 participants (58.0% male, 39.7% female, 1.8% non-binary, 0.5% preferred not to say) with a mean age of 28.7 years (SD = 9.28). Participants reported moderate familiarity with robots (M = 5.05, SD = 1.54; 1 = work/study robots daily, 7 = only know robots from popular culture) and limited healthcare-related experience (M = 5.19, SD = 2.02; 1 = very related, 7 = not related at all). Participants came from several European countries, primarily Portugal (24.7%) and Poland (23.7%), followed by Spain (8.2%), Italy (6.8%), Greece (5.5%), Hungary (3.2%), Germany (2.7%), the Czech Republic (2.7%), Estonia (1.4%), and Sweden (1.4%). Additional participants (less than 2 per country) were from other countries within Europe (19.6%), and one participant (0.5%) preferred not to disclose their country.

4.2 Study Design

The study followed a 4 × 3 mixed design. Mitigation Strategy was manipulated within subjects with four levels: human-like, robot-like, combined, and no recovery. Error Condition was manipulated between subjects with three levels: abrupt stop, corner navigation error, and close-distance navigation error, chosen to represent different navigation failures and levels of perceived severity (see Section 4.6). Participants were randomly assigned to one Error Condition (abrupt: $n = 74$, corner: $n = 73$, distance: $n = 72$) and evaluated the Baseline Condition and all four Recovery Strategies in randomized order. The primary dependent variables were *Trust* and *Perceived Social Intelligence* (PSI) (see Section 4.4).

4.3 Stimuli

We used Unity to create animations depicting robot navigation scenarios and error mitigation strategies[1].

4.3.1 Baseline Videos. Participants first viewed two baseline animations showing the robot behaving correctly. The first introduced the healthcare delivery robot's typical behavior and social capabilities. The second depicted one of three correct navigation scenarios (avoiding abrupt stops, corner surprises, or close proximity to a VH). These videos established expected robot behavior before introducing errors. A brief description of the robot's role in the hospital was shown before each video.

Social Baseline Video (14 s): The robot approaches a VH, acknowledges a greeting through movement, sound, and LED cues, and continues along its path.

Correct Navigation Videos (13 s): Correct Abrupt Stop: The robot moves aside to allow a VH to pass. *Correct Corner Navigation:* The robot turns a corner with a wide arc to avoid a VH. *Correct Distance:* The robot navigates around a VH taking a phone call.

4.3.2 Error Videos. Three videos (8 s) depicted navigation errors (Figure 1). Before each video, participants were informed that the robot would make an error. *Abrupt Stop Error:* The robot abruptly stops due to an obstacle, nearly causing a collision with a VH behind it. *Corner Error:* The robot turns a corner and nearly collides with a

[1]The complete set of video stimuli used in this study is publicly available on the Open Science Framework (OSF): https://doi.org/10.17605/OSF.IO/BA9K5



Figure 1: Screen capture of video animations showcasing the robot error. Left to right: abrupt, corner, and distance errors.



Figure 2: Screen capture of animated video showcasing the human-like recovery strategy.

VH. *Distance Error:* The robot drives too closely past a VH talking on the phone.

4.3.3 Error and Recovery Videos. These videos began with one of the error scenarios, after which the camera focused on the robot demonstrating a mitigation strategy. Video length varied depending on the strategy. *Human-like recovery (23 s):* This response constitutes a non-verbal expression of regret, without explicit verbal or textual content. The robot turns toward the VH, displays sad eyes, changes its LED color to purple, emits an apologetic sound, and then resumes its path (Figure 2). *Robot-like recovery (31 s):* The robot turns toward the VH, changes LEDs to yellow, and displays an error message followed by a message indicating that a report has been generated to prevent the issue from recurring (Figure 3). *Human-like + Robot-like recovery (33 s):* The robot first performs the apology behavior and then displays the error and learning messages. *None (14 s):* The robot makes the error but does not acknowledge it and continues its path.

4.4 Measures

Multi-Dimensional Measure of Trust (MDMT): We used an adapted version of the MDMT capability subscale (capable, skilled, competent), which measures perceived capability and trustworthiness [43]. To improve comprehension among our multilingual sample, the item meticulous was replaced with intelligent, and trustworthy was added to capture participants' direct assessment of the robot's trustworthiness.

Perceived Social Intelligence (PSI): We used the Social Competence subscale of the PSI scale [1], which assesses perceived social competence (e.g., socially competent, socially aware, socially clueless, and having strong social skills).

All items were rated on 7-point Likert-type scales ranging from 1 (strongly disagree) to 7 (strongly agree). Items from both scales were presented in a single randomized list to minimize order effects.

4.5 Procedure

The study was conducted online using Qualtrics and distributed through Prolific. After providing informed consent, participants watched two baseline videos showing a healthcare delivery robot interacting with a virtual human (VH) in a hospital corridor: one demonstrating the robot's social behavior and one showing correct navigation. Participants then described what they observed and completed measures of *Trust* (MDMT) and *PSI*. Participants next watched a video depicting a navigation error, described the event, rated its severity on a 7-point scale, and again completed the *Trust* and *PSI* measures. Finally, participants viewed four videos showing the error followed by different mitigation strategies. The order of these videos was randomized. After each video, participants described the scenario and completed the same measures.

At the end of the study, participants provided demographic information including age, gender, country of residence, and familiarity with robots.

4.6 Analysis

Analyses were conducted in R using linear mixed-effects models (LMMs) fitted with the `lme4` and `lmerTest` packages. Separate models were fitted for Trust and PSI recovery. Error condition and recovery strategy were included as fixed effects and participant ID as a random effect. Additional predictors included initial loss, age, gender, healthcare experience, and robot familiarity. Initial loss was calculated as the difference between pre- and post-error ratings and included to account for variation in the magnitude of the initial decline. Model selection followed a hierarchical BIC-based procedure, progressively adding predictors and interaction terms [19]. Internal consistency was high for both the Trust ($\alpha = .94$) and PSI ($\alpha = .88$) scales.

A pilot study confirmed that the three navigation errors differed in perceived severity ($F(2, 216) = 18.57$, $p < .001$, $\eta^2 = .147$). Error condition therefore served as the operationalization of perceived severity and perceived severity was not included separately in the models.

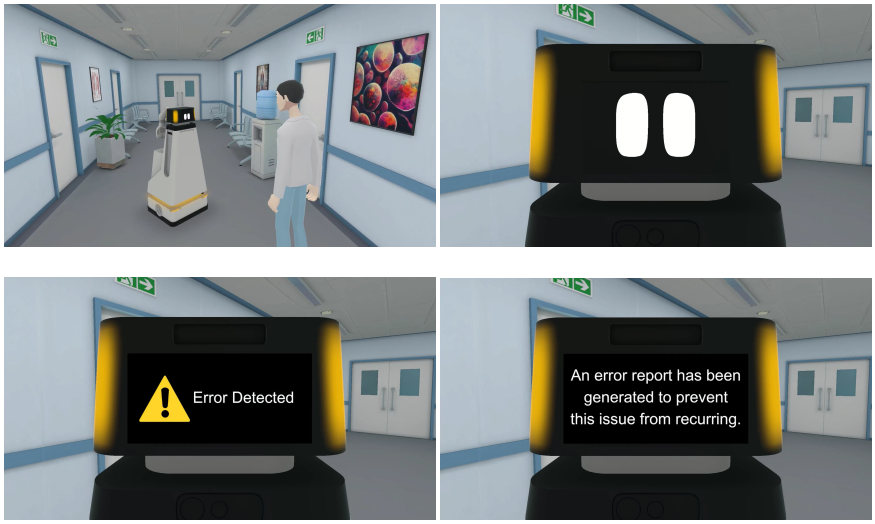


Figure 3: Screen capture of video animations showcasing the robot-like recovery strategy.

5 Results

5.1 Effects of Errors on Trust and PSI

To assess whether participants' trust and perceptions of the robot decreased after a navigation error **(H1a)**, we conducted paired-samples t-tests comparing ratings in the baseline condition with ratings after the robot made an error. The results showed that both *Trust* and *Perceived Social Intelligence (PSI)* were significantly lower after the robot made an error than in the baseline condition. *Trust* decreased by 1.55 points, $t(218) = -19.07$, $p < .001$, 95% CI [−1.71, −1.39]. Similarly, *PSI* decreased by 1.14 points, $t(218) = -13.51$, $p < .001$, 95% CI [−1.30, −0.97]. A paired-samples t-test comparing the magnitude of the decreases showed that *Trust* dropped significantly more than *PSI*, $t(218) = 4.64$, $p < .001$.

To examine whether the perceived *Severity* of an error was associated with the loss of *Trust* and *PSI* **(H1b)**, we conducted Pearson correlation analyses. The results revealed a moderate positive correlation between perceived error *Severity* and the loss of *Trust*, $r(217) = .35$, $p < .001$, 95% CI [0.23, 0.46], indicating that more severe errors were associated with larger decreases in trust. In contrast, the relationship between perceived error *Severity* and the loss of *Perceived Social Intelligence* was weak, $r(217) = .14$, $p = .035$, 95% CI [0.01, 0.27], suggesting that severity had only a small effect on participants' perceptions of the robot's perceived social intelligence.

5.2 Linear Mixed-Effects Models

Linear mixed-effects models were used to examine how recovery strategies, error conditions, and initial loss of *Trust* and *PSI* (i.e., the change from baseline to post-error) influenced post-error changes in *Trust* and *PSI*. These analyses address **H2a**, which predicts that recovery strategies influence outcomes, **H2b**, which predicts that larger initial losses affect these outcomes, and **H3a**, which predicts that participant characteristics play a role. Because the three navigation errors differed significantly in perceived severity, error type was used as an operationalization of severity when testing whether recovery effectiveness varied across errors (**H3b**).

5.2.1 Trust. Descriptive statistics (Table 6 in Appendix B) provide an overview of differences in *Trust* across error types and recovery strategies (see Figure 4).

Model fit statistics for the linear mixed-effects models are presented in Table 3 in Appendix A. Models were built hierarchically by adding predictors and interactions. Additional co-variates (gender, healthcare experience, and robot familiarity) did not improve model fit. Although age showed a marginal improvement ($\chi^2(1) = 3.90$, $p = .048$), the increase in model complexity was not supported by the BIC. Model 7, which included interactions between recovery strategy, error condition, and initial loss of *Trust*, provided the best fit ($\chi^2 = 88.67$, $p < .001$) and was therefore selected for further analysis.

The results of the linear mixed-effects model (Model 7 in Table 3) analyzing trust as an outcome are presented in Table 1.

Effect of Recovery Strategy. Compared to the reference condition (apology and learning combined), *Trust* was significantly lower when no recovery was provided ($\beta = -0.81$, $SE = 0.14$, $t(645) = -5.78$, $p < .001$). Learning alone also reduced significantly *Trust* ($\beta = -0.30$, $SE = 0.14$, $p = .032$), while apology alone showed no significant effect ($p = .061$).

Effect of Error Condition. Error type did not significantly influence *Trust*, as neither corner nor distance errors differed significantly from abrupt errors.

Effect of Initial Loss of Trust. Initial loss of *Trust* strongly predicted *Trust* recovery ($\beta = 0.77$, $SE = 0.05$, $t = 14.58$, $p < .001$).

Interaction Effects. Interactions between recovery strategy and error condition were not significant. However, the interaction between recovery strategy and initial loss of *Trust* was significant for apology ($\beta = -0.15$, $p = .008$) and for no recovery ($\beta = -0.51$, $p < .001$), indicating reduced recovery when trust loss was larger.

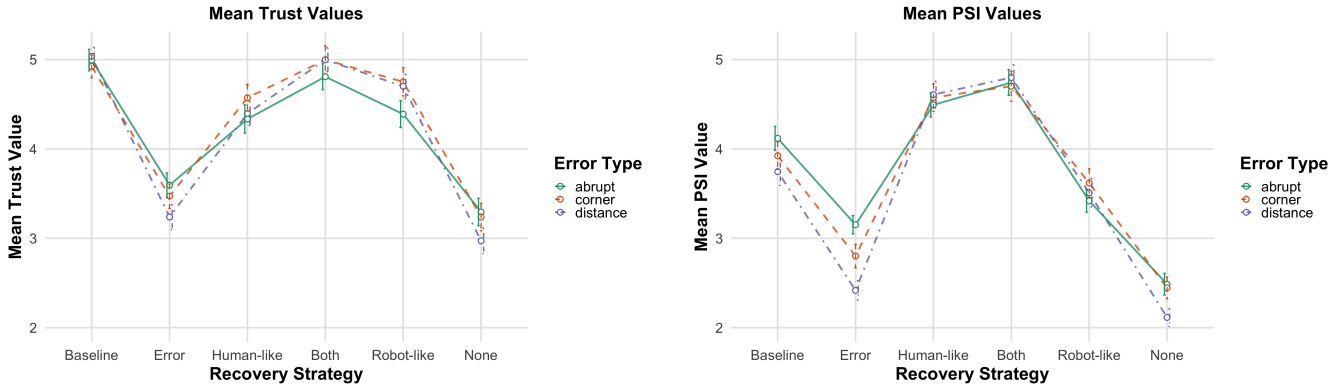


Figure 4: Averaged *Trust* (left) and *Perceived Social Intelligence (PSI)* (right) before and after the error, and across recovery strategies. *Trust* decreases after the error, remains lowest without recovery, and recovers most in the combined condition. *PSI* shows a similar pattern, with human-like recovery elevating ratings above baseline and robot-like recovery remaining below baseline.

Table 1: Trust: Recovery Strategy, Error Condition, and Their Interaction. Significance levels: * p < .05, ** p < .01, *** p < .001

Effect	Estimate (b)	Std. Error (SE)	t-value	p-value	95% CI
Intercept	0.14	0.13	1.04	.297	[-0.38, 0.13]
Recovery: Human-like	-0.26	0.14	-1.87	.061	[-0.31, 0.23]
Recovery: Robot-like	-0.30	0.14	-2.16	**.032***	[-0.81, -0.27]
Recovery: None	-0.81	0.14	-5.78	**<.001*****	[-0.82, -0.27]
Error: Corner	0.27	0.15	1.78	.075	[0.03, 0.63]
Error: Distance	0.24	0.15	1.55	.123	[-0.13, 0.47]
Initial Loss of Trust	0.77	0.05	14.58	**<.001*****	[0.52, 0.72]
Recovery: Human-like x Error: Corner	0.06	0.16	0.34	.731	[-0.37, 0.26]
Recovery: Robot-like x Error: Corner	0.18	0.16	1.08	.279	[-0.20, 0.44]
Recovery: None x Error: Corner	-0.22	0.16	-1.34	.180	[-0.59, 0.04]
Recovery: Human-like x Error: Distance	-0.07	0.17	-0.41	.685	[-0.25, 0.39]
Recovery: Robot-like x Error: Distance	0.16	0.17	0.94	.348	[-0.10, 0.54]
Recovery: None x Error: Distance	-0.31	0.17	-1.88	.060	[-0.57, 0.08]
Recovery: Human-like x Initial Loss of Trust	-0.15	0.06	-2.68	**.008****	[0.04, 0.26]
Recovery: Robot-like x Initial Loss of Trust	-0.08	0.06	-1.50	.134	[-0.04, 0.18]
Recovery: None x Initial Loss of Trust	-0.51	0.06	-8.97	**<.001*****	[-0.46, -0.24]

Note. $R^2_{\text{marg.}} = .544$; $R^2_{\text{cond.}} = .740$.

Post-hoc Comparisons. Tukey-adjusted post-hoc comparisons within each error condition showed that the combined human-like and robot-like strategy resulted in significantly higher *Trust* than the human-like strategy across all error types (all $p < .001$), and higher than the robot-like strategy in the abrupt condition ($p = .0011$), but not in the corner ($p = .1248$) or distance conditions ($p = .0876$). The human-like and robot-like strategies did not significantly differ in any condition. Across all error types, both strategies yielded significantly higher *Trust* than the no-recovery condition (all $p < .001$).

5.2.2 Perceived Social Intelligence. Model fit statistics for the PSI models are presented in Table 4 in Appendix A. Models were built hierarchically by adding predictors and interactions. Adding error condition significantly improved model fit ($\chi^2 = 18.31$, $p < .001$), and including initial loss of *PSI* further improved the model ($\chi^2 = 70.39$, $p < .001$). Additional co-variates (gender, healthcare experience, and robot familiarity) and interaction terms did not improve model fit and increased AIC/BIC values. Therefore, Model 3, which included recovery strategy, error condition, and initial loss of *PSI*, was selected as the final model.

The results of the linear mixed-effects model analyzing PSI as an outcome are presented in Table 2.

Effect of Recovery Strategy. Compared to the reference condition (apology and learning combined), *PSI* was significantly lower when only apology was provided ($\beta = -0.19$, $SE = 0.09$, $p = .041$). Learning alone led to a larger decrease in *PSI* ($\beta = -1.23$, $SE = 0.09$, $p < .001$), while no recovery produced the strongest reduction ($\beta = -2.40$, $SE = 0.09$, $p < .001$).

Table 2: PSI: Recovery Strategy, Error Condition, and PSI Delta. Significance levels: * p < .05, ** p < .01, *** p < .001

Effect	Estimate (b)	Std. Error (SE)	t-value	p-value	95% CI
Intercept	1.185	0.1218	9.74	**< .001*** **	[0.88, 1.49]
Recovery: Human-like	-0.189	0.0925	-2.05	**.041***	[-0.38, -0.01]
Recovery: Robot-like	-1.232	0.0925	-13.32	**< .001*** **	[-1.42, -1.05]
Recovery: None	-2.397	0.0925	-25.93	**< .001*** **	[-2.58, -2.16]
Error: Corner	0.335	0.1394	2.41	**.017***	[0.03, 0.63]
Error: Distance	0.556	0.1407	3.95	**< .001*** **	[0.28, 0.83]
Initial Loss of PSI	0.417	0.0462	9.03	**< .001*** **	[0.33, 0.51]

Note. $R^2_{\text{marg.}} = .472$; $R^2_{\text{cond.}} = .650$.

Effect of Error Condition. Error type significantly influenced *PSI*. Compared to abrupt errors, both corner errors ($\beta = 0.34$, $SE = 0.14$, $p = .017$) and distance errors ($\beta = 0.56$, $SE = 0.14$, $p < .001$) were associated with higher *PSI* ratings.

Effect of Initial Loss of PSI. Initial loss of *PSI* significantly predicted recovery in *PSI* ($\beta = 0.42$, $SE = 0.046$, $p < .001$), indicating that larger initial declines were associated with stronger recovery after recovery.

Post-hoc Comparisons. Tukey-adjusted comparisons showed that *PSI* was higher for the combined human-like and robot-like strategy and the human-like strategy, which did not differ from each other ($p = .171$), but both exceeded the robot-like strategy and no recovery (all $p < .001$). The robot-like strategy also resulted in higher *PSI* than no recovery ($p < .001$).

6 Discussion

This study examined how navigation errors and different recovery strategies (human-like, robot-like, their combination, and no recovery strategy) influence users' trust and perceptions of a robot's social intelligence. In this section, we will discuss the results, hypotheses, and implications for the field of human-robot interaction.

6.1 Effects on Trust

As expected, *trust in the robot decreased significantly following a navigation error* (H1a). This aligns with prior work showing that robot errors undermine perceived reliability and competence [25, 28, 34, 36, 41]. Notably, the decrease in trust was larger than the decrease in perceived social intelligence, suggesting that navigation errors are primarily interpreted as competence-related failures rather than social ones [40]. One possible explanation is that in a healthcare context, users prioritize safety and reliability, making functional failures more salient than social ones.

We further expected that more severe errors would lead to larger decreases in trust (H1b), which was supported. *Perceived severity was moderately correlated with trust loss, indicating that more severe navigation failures produce stronger negative reactions.* This finding aligns with prior work showing that the perceived severity of robot errors can influence users' responses and evaluations of robot performance [9, 20]

Recovery strategies played a central role in restoring trust. While we expected robot-like strategies to be most effective (H2a), results indicate that *the robot-like strategy was at least as effective as the human-like strategy*, highlighting the importance of competence-based signals. *The strongest trust recovery was observed when both strategies were combined.* This aligns with prior work showing that combining expressive and informative behaviors supports effective recovery [25]. It also suggests that addressing the functional consequences of an error may be sufficient to restore trust in contexts where social expectations toward the robot are relatively low [16]. However, integrating social cues appears to provide a more complete response.

Contrary to expectations (H2b) and prior work suggesting that the effectiveness of recovery strategies depends on the severity of the failure [4], *larger initial decreases in trust were associated with greater recovery following recovery.* However, this pattern depended on the type of recovery. Human-like responses alone became less effective at higher levels of trust loss, and the absence of mitigation strategy was particularly detrimental when initial trust loss was substantial. In contrast, the combined human-like and robot-like strategy remained effective across levels of trust loss. Because both variables are derived from related difference scores, this relationship should be interpreted cautiously.

6.2 Effects on Social Intelligence

Perceived social intelligence also decreased following an error (H1a), although less than trust. This suggests that navigation failures have a more limited effect on users' perceptions of a robot's social abilities than on evaluations of its competence[40]. One possible explanation is that users may maintain relatively stable expectations about a robot's social capabilities, even when its performance fails. Consistent with this interpretation, *perceived error severity was only weakly correlated with changes in PSI* (H1b), one possible explanation is that severity had only a limited influence on perceptions of the robot's social intelligence. This contrasts with trust, where severity played a clearer role, further supporting the idea that functional and social evaluations are reshaped differently.

Although we expected human-like strategies to have the strongest effect on PSI (H2a), *PSI ratings were highest when human-like and robot-like responses were combined, followed by human-like responses alone, and lowest for robot-like responses.* These results are consistent with prior work showing that anthropomorphic behaviors enhance perceived social capability [36] and that combining expressive and informative behaviors supports effective recovery [25]. Notably, *human-like responses increased PSI beyond baseline levels* (see Figure 4). This pattern may indicate that such behaviors lead users to perceive the robot as more socially capable than before the error occurred. One possible interpretation is that anthropomorphic

cues signal social awareness and intentionality, leading users to attribute broader socio-affective abilities to the robot. While this may support user acceptance [23, 45], it also raises important design considerations. PSI reflects perceived social capability, but not necessarily the robot's actual abilities. Such increases may contribute to mismatches between what users expect and what the robot can deliver [27, 42].

Finally, similar to trust and contrary to expectations based on prior work [4], *larger initial decreases in PSI were associated with greater recovery* (H2b), suggesting that more noticeable errors create stronger opportunities for recovery strategies to influence user perceptions. This drop allows social evaluations to be not only restored but, in some cases, enhanced. At the same time, this relationship should be interpreted with caution, as both the initial loss and recovery measures are derived from related difference scores.

6.3 Role of User and Error Characteristics

We expected that participant characteristics would influence the effectiveness of recovery strategies for trust and PSI (H3a), in line with prior work suggesting that user characteristics can shape perceptions of robots and human–robot interaction [6, 14, 26]. However, the results did not support this hypothesis. *None of the examined participant characteristics significantly improved the models predicting trust or perceived social intelligence,* indicating that participants responded to the robot's error recovery strategies in relatively similar ways regardless of their background (i.e., gender and age), prior experience with robots, or experience in the field of healthcare. This suggests that the observed effects may be relatively robust across different user groups.

We also expected that error severity would influence recovery effectiveness (H3b). Contrary to expectations [4, 9, 48], the results provided limited support for this hypothesis. Although the three navigation errors differed in perceived severity, and higher severity was associated with greater initial loss of trust, the interaction between error type and recovery strategy was not significant in the trust model, and mitigation effectiveness did not substantially vary across error types for perceived social intelligence. Some differences in trust across error conditions were observed in post-hoc comparisons; however, these exploratory patterns were not supported by significant interaction effects in the model. Overall, *the choice of recovery strategy appeared to play a more influential role than the severity of the error itself.*

6.4 Design Implications

These findings suggest that recovery strategies should be selected according to the desired outcome. Robot-like responses effectively restore trust while avoiding increases in perceived social intelligence. Human-like responses enhance perceptions of social intelligence but may also elevate users' expectations of the robot's capabilities. When both outcomes are desirable, combining functional and socio-emotional responses appears to provide the most effective recovery. For service robots, particularly in healthcare where reliability is critical, recovery strategies should be carefully calibrated [42]. Rather than relying solely on anthropomorphic behaviors, designers should use competence-based signals as the foundation of recovery and incorporate social cues only where appropriate. Recovery behaviors should remain aligned with the robot's actual capabilities to support ethical design and avoid creating unrealistic user expectations [27, 32].

7 Limitations and Future Work

This study has several limitations. First, the experiment was conducted using an online video-based methodology with animated virtual interactions, meaning participants observed robot behavior rather than interacting with the robot directly. Consequently, the findings may not fully generalize to in-person interactions with physical robots, where embodiment, motion, and other contextual factors may influence users' perceptions. Second, the recovery strategies differed in duration and presentation, which may have influenced participants' evaluations independently of the intended recovery behavior. Furthermore, we did not explicitly validate whether participants perceived the recovery behaviors as intended (i.e., as robot-like or human-like). Future work should control for these factors and validate the intended manipulation of recovery strategy. Third, the findings are based on a single robot platform, healthcare scenario, navigation task, and set of recovery behaviors, as well as a participant sample primarily from Europe. These factors may limit generalizability across robot embodiments, application domains, and cultural contexts. Future work should therefore validate these findings with physical robots in real-world deployments and over longer interactions to better understand how trust and perceived social intelligence evolve over time. Replicating the study with different robots, tasks, and environments would also help assess the robustness of the findings. Finally, future research should further investigate how the magnitude of initial trust and perceived social intelligence loss influences the effectiveness of recovery strategies.

8 Conclusion

This study examined how robots can recover from navigation errors in human–robot interaction, particularly in sensitive settings such as hospitals. We found that while errors reduce both Trust and Perceived Social Intelligence (PSI), these effects can be mitigated through appropriate recovery strategies. Both human-like and robot-like responses supported trust recovery, but shaped social perceptions in different ways. Human-like responses elevated perceived social intelligence beyond baseline levels, whereas robot-like responses restored trust without substantially altering social expectations. The strongest recovery was observed when both strategies were combined. These findings suggest a key design implication for service robots: recovery strategies should be carefully calibrated to the robot's role, combining functional signals of competence with social cues where appropriate, while avoiding over-reliance on anthropomorphic behaviors that may create mismatches between perceived and actual capabilities.

Acknowledgments

This work was made possible through funding from the European Union's Horizon 2020 research and innovation program under grant agreement no: 101017008 (HARMONY) and from the Netherlands Organisation for Scientific Research, IMEC-One Planet and other private parties, under grant agreement no: 17626 (Synergia).

References

[1] Kimberly A Barchard, Leiszle Lapping-Carr, R Shane Westfall, Andrea Fink-Armold, Santosh Balajee Banisetty, and David Feil-Seifer. 2020. Measuring the perceived social intelligence of robots. *ACM Transactions on Human-Robot Interaction (THRI)* 9, 4 (2020), 1–29.

[2] M. Biswas and J.C. Murray. 2015. Towards an imperfect robot for long-term companionship: case studies using cognitive biases. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Hamburg, Germany, 5978–5983. doi:10.1109/IROS.2015.7354228

[3] British Standards Institute. 2016. *BS 8611:2016 Robots and Robotic Devices: Guide to the Ethical Design and Application of Robots and Robotic Systems*. Technical Report. British Standards Institute.

[4] Daniel J. Brooks, Momotaz Begum, and Holly A. Yanco. 2016. Analysis of reactions towards failures and recovery strategies for autonomous robots. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, New York, NY, USA, 487–492. doi:10.1109/ROMAN.2016.7745162

[5] Dito Eka Cahya and Manuel Giuliani. 2021. Appropriate Robot Reactions to Erroneous Situations in Human-Robot Collaboration. In *Social Robotics*, Haizhou Li, Shuzhi Sam Ge, Yan Wu, Agnieszka Wykowska, Hongsheng He, Xiaorui Liu, Dongyu Li, and Jairo Perez-Osorio (Eds.). Vol. 13086. Springer International Publishing, Cham, 166–177. Series Title: Lecture Notes in Computer Science. doi:10.1007/978-3-030-90525-5_15

[6] Dito Eka Cahya and Manuel Giuliani. 2021. Appropriate Robot Reactions to Erroneous Situations in Human-Robot Collaboration. In *Social Robotics: 13th International Conference, ICSR 2021, Singapore, Singapore, November 10–13, 2021, Proceedings 13*. Springer, 166–177.

[7] David Cameron, Stevienna De Saille, Emily C. Collins, Jonathan M. Aitken, Hugo Cheung, Adriel Chua, Ee Jing Loh, and James Law. 2021. The effect of social-cognitive recovery strategies on likability, capability and trust in social robots. *Computers in Human Behavior* 114 (Jan. 2021), 106561. doi:10.1016/j.chb.2020.106561

[8] Sungwoo Choi, Anna S. Mattila, and Lisa E. Bolton. 2021. To Err Is Human(-oid): How Do Consumers React to Robot Service Failure and Recovery? *Journal of Service Research* 24, 3 (Aug. 2021), 354–371. doi:10.1177/1094670520978798

[9] Filipa Correia, Carla Guerra, Samuel Mascarenhas, Francisco S Melo, and Ana Paiva. 2018. Exploring the impact of fault justification in human-robot trust. In *Proceedings of the 17th international conference on autonomous agents and multiagent systems*. 507–513.

[10] Ewart J. De Visser, Samuel S. Monfort, Ryan McKendrick, Melissa A. B. Smith, Patrick E. McKnight, Frank Krueger, and Raja Parasuraman. 2016. Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied* 22, 3 (Sept. 2016), 331–349. doi:10.1037/xap0000092

[11] Connor Esterwood and Lionel P. Robert Jr. 2023. Three Strikes and you are out!: The impacts of multiple human–robot trust violations and repairs on robot trustworthiness. *Computers in Human Behavior* 142 (May 2023), 107658. doi:10.1016/j.chb.2023.107658

[12] Connor Esterwood and Lionel Robert. 2023. The theory of mind and human–robot trust repair. *Scientific Reports* 13 (06 2023). doi:10.1038/s41598-023-37032-0

[13] Connor Esterwood and Lionel P. Robert. 2021. Do You Still Trust Me? Human-Robot Trust Repair Strategies. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, Vancouver, BC, Canada, 183–188. doi:10.1109/RO-MAN50785.2021.9515365

[14] Connor Esterwood and Lionel P. Robert. 2022. Having the Right Attitude: How Attitude Impacts Trust Repair in Human—Robot Interaction. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Sapporo, Japan, 332–341. doi:10.1109/HRI53351.2022.9889535

[15] Connor Esterwood and Lionel P. Robert. 2022. A Literature Review of Trust Repair in HRI. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Napoli, Italy, 1641–1646. doi:10.1109/RO-MAN53752.2022.9900667

[16] Connor Esterwood and Lionel P. Robert. 2025. Repairing Trust in Robots?: A Meta-analysis of HRI Trust Repair Studies with A No-Repair Condition :: Data ::. doi:10.5281/ZENODO.14583133

[17] Helena Anna Frijns, Oliver Schürer, and Sabine Theresia Koeszegi. 2023. Communication Models in Human–Robot Interaction: An Asymmetric MODel of ALterity in Human–Robot Interaction (AMODAL-HRI). *International Journal of Social Robotics* 15, 3 (March 2023), 473–500. doi:10.1007/s12369-021-00785-7

[18] Hideki Garcia Goo, Bob R. Schadenberg, Jan Kolkmeier, Khiet P. Truong, and Vanessa Evers. 2025. Being Sorry is the Hardest Thing: How Robots can Apologize and Learn from Mistakes to Restore People's Trust. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–9. doi:10.1145/3706599.3719739

[19] Mark Gavin. 2004. Hierarchical linear models: Applications and data analysis methods. *Organizational Research Methods* 7, 2 (2004), 228.

[20] Romi Gideoni, Shanee Honig, and Tal Oron-Gilad. 2024. Is It Personal? The Impact of Personally Relevant Robotic Failures (PeRFs) on Humans' Trust, Likeability, and Willingness to Use the Robot. *International Journal of Social Robotics* 16, 6 (June 2024), 1049–1067. doi:10.1007/s12369-022-00912-y

[21] Manuel Giuliani, Nicole Mirnig, Gerald Stollnberger, Susanne Stadler, Roland Buchner, and Manfred Tscheligi. 2015. Systematic analysis of video data from different human–robot interaction studies: a categorization of social signals during error situations. *Frontiers in Psychology* 6 (July 2015). doi:10.3389/fpsyg.2015.00931

[22] Kasper Hald, Katharina Weitz, Elisabeth André, and Matthias Rehm. 2021. "An Error Occurred!" - Trust Repair With Virtual Robot Using Levels of Mistake Explanation. In *Proceedings of the 9th International Conference on Human-Agent Interaction*. ACM, Virtual Event Japan, 218–226. doi:10.1145/3472307.3484170

[23] Adriana Hamacher, Nadia Bianchi-Berthouze, Anthony G. Pipe, and Kerstin Eder. 2016. Believing in BERT: Using expressive communication to enhance trust and counteract operational error in physical Human-robot interaction. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, New York, NY, USA, 493–500. doi:10.1109/ROMAN.2016.7745163

[24] Shanee Honig and Tal Oron-Gilad. 2018. Understanding and Resolving Failures in Human-Robot Interaction: Literature Review and Model Development. *Frontiers in Psychology* 9 (June 2018), 861. doi:10.3389/fpsyg.2018.00861

[25] E. S. Kox, J. H. Kerstholt, T. F. Hueting, and P. W. de Vries. 2021. Trust repair in human-agent teams: the effectiveness of explanations and expressing regret. *Autonomous Agents and Multi-Agent Systems* 35, 2 (Oct. 2021), 30. doi:10.1007/s10458-021-09515-9

[26] Johannes Maria Kraus, Julia Merger, Felix Gröner, and Jessica Pätz. 2023. 'Sorry' Says the Robot: The Tendency to Anthropomorphize and Technology Affinity Affect Trust in Repair Strategies after Error. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Stockholm Sweden, 436–441. doi:10.1145/3568294.3580122

[27] Makoto Kureha. 2023. On the moral permissibility of robot apologies. *AI & SOCIETY* (Sept. 2023). doi:10.1007/s00146-023-01782-2

[28] Min Kyung Lee, Sara Kiesler, Jodi Forlizzi, Siddhartha Srinivasa, and Paul Rybski. 2010. Gracefully mitigating breakdowns in robotic services. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Osaka, Japan, 203–210. doi:10.1109/HRI.2010.5453195

[29] Christoforos Mavrogiannis, Francesca Baldini, Allan Wang, Dapeng Zhao, Pete Trautman, Aaron Steinfeld, and Jean Oh. 2023. Core Challenges of Social Robot Navigation: A Survey. *ACM Transactions on Human-Robot Interaction* 12, 3 (Sept. 2023), 1–39. doi:10.1145/3583741

[30] Nicole Mirnig, Gerald Stollnberger, Markus Miksch, Susanne Stadler, Manuel Giuliani, and Manfred Tscheligi. 2017. To Err Is Robot: How Humans Assess and Act toward an Erroneous Social Robot. *Frontiers in Robotics and AI* 4 (May 2017), 21. doi:10.3389/frobt.2017.00021

[31] Babiche L. Pompe, Ella Velner, and Khiet P. Truong. 2022. The Robot That Showed Remorse: Repairing Trust with a Genuine Apology. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Napoli, Italy, 260–265. doi:10.1109/RO-MAN53752.2022.9900860

[32] Laurel D Riek and Don Howard. [n. d.]. A Code of Ethics for the Human-Robot Interaction Profession. ([n. d.]).

[33] Paul Robinette, Ayanna M Howard, and Alan R Wagner. 2015. Timing is key for robot trust repair. In *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings 7*. Springer, 574–583.

[34] Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L Walters. 2017. How the timing and magnitude of robot errors influence peoples' trust of robots in an emergency scenario. In *Social Robotics: 9th International Conference, ICSR 2017, Tsukuba, Japan, November 22-24, 2017, Proceedings 9*. Springer, 42–52.

[35] Maha Salem, Friederike Eyssel, Katharina Rohlfing, Stefan Kopp, and Frank Joublin. 2013. To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability. *International Journal of Social Robotics* 5 (05 2013), 313–323. doi:10.1007/s12369-013-0196-9

[36] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Portland Oregon USA, 141–148. doi:10.1145/2696454.2696497

[37] Satragni Sarkar, Dejanira Araiza-Illan, and Kerstin Eder. 2017. Effects of Faults, Experience, and Personality on Trust in a Robot Co-Worker. arXiv:1703.02335 [cs]. http://arxiv.org/abs/1703.02335

[38] Sarah Strohkorb Sebo, Priyanka Krishnamurthi, and Brian Scassellati. 2019. "I Don't Believe You": Investigating the Effects of Robot Trust Violation and Repair. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Daegu, Korea (South), 57–65. doi:10.1109/HRI.2019.8673169

[39] Maia Stiber and Chien-Ming Huang. 2020. Not All Errors Are Created Equal: Exploring Human Responses to Robot Errors with Varying Severity. In *Companion Publication of the 2020 International Conference on Multimodal Interaction*. ACM, Virtual Event Netherlands, 97–101. doi:10.1145/3395035.3425245

[40] Leimin Tian and Sharon Oviatt. 2021. A Taxonomy of Social Errors in Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction* 10, 2 (June

2021), 1–32. doi:10.1145/3439720

[41] Suzanne Tolmeijer, Astrid Weiss, Marc Hanheide, Felix Lindner, Thomas M. Powers, Clare Dixon, and Myrthe L. Tielman. 2020. Taxonomy of Trust-Relevant Failures and Mitigation Strategies. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Cambridge United Kingdom, 3–12. doi:10.1145/3319502.3374793

[42] Ilaria Torre and Laurence White. 2021. Trust in Vocal Human–Robot Interaction: Implications for Robot Voice Design. In *Voice Attractiveness*, Benjamin Weiss, Jürgen Trouvain, Melissa Barkat-Defradas, and John J. Ohala (Eds.). Springer Singapore, Singapore, 299–316. Series Title: Prosody, Phonology and Phonetics. doi:10.1007/978-981-15-6627-1_16

[43] Daniel Ullman and Bertram F Malle. 2019. MDMT: multi-dimensional measure of trust.

[44] Jered Vroon, Gwenn Englebienne, and Vanessa Evers. 2016. Responsive Social Agents: Feedback-Sensitive Behavior Generation for Social Interactions. In *Social Robotics*, Arvin Agah, John-John Cabibihan, Ayanna M. Howard, Miguel A. Salichs, and Hongsheng He (Eds.). Vol. 9979. Springer International Publishing, Cham, 126–137. Series Title: Lecture Notes in Computer Science. doi:10.1007/978-3-319-47437-3_13

[45] Katie Winkle, Praminda Caleb-Solly, Ute Leonards, Ailie Turton, and Paul Bremner. 2021. Assessing and Addressing Ethical Risk from Anthropomorphism and Deception in Socially Assistive Robots. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Boulder CO USA, 101–109. doi:10.1145/3434073.3444666

[46] Jin Xu and Ayanna Howard. 2022. Evaluating the Impact of Emotional Apology on Human-Robot Trust. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Napoli, Italy, 1655–1661. doi:10.1109/RO-MAN53752.2022.9900518

[47] Kai Chi Yam, Yochanan E. Bigman, Pok Man Tang, Remus Ilies, David De Cremer, Harold Soh, and Kurt Gray. 2021. Robots at work: People prefer—and forgive—service robots with perceived feelings. *Journal of Applied Psychology* 106, 10 (Oct. 2021), 1557–1572. doi:10.1037/apl0000834

[48] Hongyan Yang, Hong Xu, Yan Zhang, Yan Liang, and Ting Lyu. 2022. Exploring the effect of humor in robot failure. *Annals of Tourism Research* 95 (July 2022), 103425. doi:10.1016/j.annals.2022.103425

[49] Ceciley (Xinyi) Zhang and Sun Kyong Lee. 2026. "But You Shouldn't Blame Me": A Cross-National Comparison of the Effects of Performance Failures and Trust Repairs in Human–Robot Interactions. *ACM Transactions on Human-Robot Interaction* 15, 1 (Jan. 2026), 1–28. doi:10.1145/3746461

A Model Building

A.1 Trust Model Building

Shown in Table 3.

Table 3: ANOVA sequential comparisons for Trust model building. Significance levels: * p < .05, ** p < .01, *** p < .001

Model	df	AIC	BIC	Log-Likelihood	χ^2	p
Basic models						
(1) Recovery Only	6	2422.30	2451.00	-1205.20	-	-
(2) Recovery + Error	8	2419.30	2457.60	-1201.70	6.96	.030*
(3) Recovery + Error + Initial Loss of Trust	9	2270.10	2313.10	-1126.00	151.27	<.001***
(4) Recovery + Error + I Loss of Trust + Age	10	2268.20	2315.90	-1124.10	3.90	.048*
Interaction models						
(5) Recovery × Error	14	2415.50	2482.40	-1193.80	-	-
(6) Recovery × Error + Initial Loss of Trust	15	2266.20	2337.90	-1118.10	151.27	<.001***
(7) Mit. × Error + Init. Loss of Trust × Mit.	18	2183.60	2269.50	-1073.80	88.67	<.001***

A.2 PSI Model Building

Shown in Table 4.

Table 4: ANOVA sequential comparisons for PSI model building. Significance levels: * p < .05, ** p < .01, *** p < .001

Model	df	AIC	BIC	Log-Likelihood	χ^2	p
Basic models						
(1) Recovery Only	6	2765.60	2794.20	−1376.80	-	-
(2) Recovery + Error	8	2751.30	2789.50	−1367.60	18.31	< .001***
(3) Recovery + Error + Initial Loss of PSI	9	2682.90	2725.90	−1332.50	70.39	< .001***

B Descriptive statistics

B.1 Severity Values for Errors

Shown in Table 5.

Table 5: Summary Severity statistics of Recovery and Error types.

Recovery	Error	Count	Mean	Median	SD	Min	Max
Error	Abrupt	74	3.76	3.00	1.70	1	7
Error	Corner	73	4.37	5.00	1.65	1	7
Error	Distance	72	5.36	6.00	1.45	2	7

B.2 Trust

Shown in Table 6.

Table 6: Summary statistics of Trust, ordered by Good, Error, and others on Trust.

Recovery	Error	Count	Mean	Median	SD	Min	Max
Good	Abrupt	74	4.99	5.20	1.04	1.00	6.80
Good	Corner	73	4.92	5.00	1.08	1.20	6.80
Good	Distance	72	5.04	5.00	0.84	2.80	7.00
Error	Abrupt	74	3.59	3.60	1.18	1.00	6.20
Error	Corner	73	3.47	3.60	1.16	1.00	6.00
Error	Distance	72	3.24	3.20	1.19	1.00	6.00
Apology	Abrupt	74	4.34	4.60	1.36	1.00	6.60
Apology	Corner	73	4.57	4.80	1.27	1.00	7.00
Apology	Distance	72	4.40	4.60	1.14	2.00	6.20
Both Apology Learning	Abrupt	74	4.81	5.00	1.25	1.00	7.00
Both Apology Learning	Corner	73	5.00	5.40	1.34	1.00	7.00
Both Apology Learning	Distance	72	5.00	5.00	1.15	2.00	7.00
Learning	Abrupt	74	4.39	4.80	1.29	1.00	6.40
Learning	Corner	73	4.75	5.00	1.34	1.00	7.00
Learning	Distance	72	4.70	5.00	1.17	1.80	7.00
None	Abrupt	74	3.29	3.40	1.32	1.00	6.20
None	Corner	73	3.24	3.20	1.31	1.00	5.80
None	Distance	72	2.97	3.00	1.22	1.00	6.00

B.3 PSI

Shown in Table 7.

Table 7: Summary statistics of Recovery Strategies and Error types, ordered by Good, Error, and others on PSI values.

Recovery	Error	Count	Mean	Median	SD	Min	Max
Good	Abrupt	74	4.12	4.25	1.16	1	6.75
Good	Corner	73	3.92	4.00	1.32	1	7
Good	Distance	72	3.75	3.75	1.32	1	6.5
Error	Abrupt	74	3.15	3.25	0.89	1	5.5
Error	Corner	73	2.80	2.75	1.13	1	6.25
Error	Distance	72	2.42	2.50	0.93	1	5
Apology	Abrupt	74	4.49	4.75	1.17	1.25	7
Apology	Corner	73	4.58	4.75	1.32	1	7
Apology	Distance	72	4.61	4.75	1.27	1	7
Both Apology Learning	Abrupt	74	4.74	5	1.24	1	7
Both Apology Learning	Corner	73	4.70	5	1.44	1	7
Both Apology Learning	Distance	72	4.80	5	1.23	1	7
Learning	Abrupt	74	3.42	3.75	1.09	1	5.75
Learning	Corner	73	3.62	3.75	1.35	1	6.25
Learning	Distance	72	3.51	3.75	1.37	1	6.25
None	Abrupt	74	2.49	2.50	1.05	1	5.25
None	Corner	73	2.45	2.50	1.04	1	5.5
None	Distance	72	2.11	2.00	0.84	1	4