This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

Having Too Many Robots for Apologies is Not Beneficial: Number and Cultural Effects of Apologizing Robots between Japan and the U.S.

Masahiro Shiomi[a]*, Taichi Hirayama[a,b], Mitsuhiko Kimoto[a,c], Takamasa Iio[a,b], and Katsunori Shimohara[a,b]

[a]*Interaction Science Laboratories, ATR, Seika-cho, Kyoto, Japan*

[b]*Doshisha University, Kyoto, Japan*

[c]*Meiji University, Kanagawa, Japan*

*Corresponding author: Masahiro Shiomi, m-shiomi@atr.jp

Masahiro Shiomi received M. Eng. and Ph.D. degrees in engineering from Osaka University in 2004 and 2007, respectively. During that same time, he was an intern researcher at the Intelligent Robotics and Communication Laboratories (IRC). He is currently a group leader in the Agent Interaction Design department at the Interaction Science Laboratories (ISL) and Interaction Technology Bank (ITB) at the Advanced Telecommunications Research Institute International (ATR). His research interests include human-robot interaction, social touch, robotics for childcare, networked robots, and field trials.

Taichi Hirayama received a M. Eng. degree in 2024 from Doshisha University, Kyoto, Japan. He was affiliated with Doshisha University at the time of contribution. His research interests include social touch, social robots, and human-robot interaction.

Mitsuhiko Kimoto received his M.E. and Ph.D. degrees from Doshisha University, Kyoto, Japan, in 2016 and 2019, respectively. He is currently a Senior Assistant Professor at the School of Science and Technology, Meiji University. From 2019 to 2022, he was a JSPS Research Fellow (PD) at the Faculty of Science and Technology, Keio University. From 2022 to 2024, he was a Research Scientist at the Interaction Science Laboratories, Advanced Telecommunications Research Institute International (ATR). His research interests include human-robot interaction, human-agent interaction, and interactive AI.

Takamasa Iio received the Ph.D. degree in engineering from Doshisha University, Kyoto, Japan, in 2012. Then, he has worked at Intelligent Robotics and Communication Laboratories, ATR, Osaka University, and the University of Tsukuba. Currently, he is an associate professor at Doshisha University, Kyoto, Japan. His field of expertise is social robotics. He is interested in

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

how people's cognition and behavior change through interaction with social robots and how human society changes.

Katsunori Shimohara received B.E. and M.E. degrees in Computer Science and Communication Engineering and a Doctor of Engineering from Kyushu University, Fukuoka, Japan in 1976, 1978, and 2000. From 2001 to 2006, he was the director of the Network Informatics Laboratories and the Human Information Science Laboratories at Advanced Telecommunications Research Institute International (ATR), Kyoto, Japan. He is a professor emeritus at Doshisha University, Kyoto, Japan. His research interests include community systems design, relationality design, and relationality-oriented systems design.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

Having Too Many Robots for Apologies is Not Beneficial: Number and Cultural Effects of Apologizing Robots between Japan and the U.S.

Effective apology behavior design has become increasingly important as social robots continue to be integrated into daily environments. Previous studies conducted in Japan suggest that when multiple robots apologize together, people tend to feel greater acceptance and trust than when only a single robot apologizes. However, it remains unclear whether 1) these findings hold in other cultural contexts and 2) how many robots are suitable for effective apologies. To answer these questions, we conducted two online surveys to investigate perceived impressions from participants in Japan and the United States. In the first survey, we used visual stimuli that depicted either one or two robots apologizing and measured the participants' forgiveness and trust toward the robot(s). The participants in both countries viewed multi-robot apologies as more acceptable than single-robot apologies, aligning with earlier findings from Japan. In the second survey, we used visual stimuli that depicted apologizing by two, four, or six robots and measured the participants' feelings of forgiveness and trust toward the robots. The participants in both countries viewed apologies from two robots as more acceptable than from either four or six robots.

Keywords: robot apology; multiple robots; robot trust; forgiveness

1. Introduction

Social robots are increasingly being deployed worldwide in diverse, daily environments, a trend that accelerated by the COVID-19 pandemic when such robots helped minimize infection risks [1, 2]. Ongoing advances in interactive functions, along with improvements in remote operation, have facilitated the use of social robots in various public and communal settings, including cafeterias [3, 4], retail spaces [3-6], mass transportation systems [7-9], exhibitions [10-12], and senior-care facilities [13-15].

To perform effectively in these settings, social robots must be perceived as trustworthy and reliable by the people with whom they interact. Accordingly, researchers have widely studied trust-building mechanisms in human-robot and human-

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

agent interactions [16-21]. Attention to ethical dimensions, including gender considerations in robot design and decision-making, has been highlighted to enhance perceived trustworthiness [22-25]. Such efforts collectively underscore the importance of ensuring that as social robots become more prevalent, they will remain both technically capable and socially competent.

Related to achieving trustworthy social robots, because even well-designed robots can make mistakes, understanding how they should apologize is critical for maintaining positive interactions. In fact, some studies have already investigated the effects of apology strategies to mitigate robots' failures [26, 27]. For instance, these studies reported that robot apologies with the intention of recovering from a failed situation are more effective than simply apologizing for the robot's mistake. Recent research has focused on using multiple robots to deliver apologies [28]. This study found that in Japan apologies offered by two robots (rather than just one) distinctly affected attitudes toward forgiveness and perceived trust.

However, the existing work on multi-robot apologies has been limited to Japan and has only investigated a maximum of two robots [28]. Because prior studies argued that cultural norms can shape the function and meaning of apologies [29-32], it is unclear whether the positive effects observed for multiple robots in a Japanese context are generalizable to other cultures as well as the appropriate number of robots for apologizing behaviors. Examining the influence of multiple-robot apologies in different cultural settings with various numbers of robots is therefore vital for developing robust design principles for social robot behaviors.

Therefore, this study addresses whether the findings from Japan apply to the United States and investigates effective robot numbers in apology settings (Fig. 1). Although this paper is an extended version of previous work by Shiomi et al. [33], we added a

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

new research question from the perspective of the effective numbers for robot apologies; accordingly, this paper contains an additional experiment, an analysis of its results, and more detailed discussions.

Figure 1. Multiple robots' apologies (captured from our experiment's visual stimuli)

2. Theoretical Frameworks

2.1 Apology and cultural differences in human-human and human-robot interaction

In human-human interaction, apologizing is a common strategy to mitigate failures not only in human-human but also human-robot interactions. In fact, several studies have already been conducted to understand the meaning of apologies and their trust-repairing processes [34-37]. In particular, the importance of apologizing with an internal or external attribution on trust-repairing has been broadly investigated [38, 39]. Different studies explored apologies as a means of reinforcing social bonds [34], highlighted their potential to strengthen relationships [35], and investigated cost-related factors for designing effective apology behaviors [36, 37]. Computational perspectives on apologies have also been examined in game-theoretic contexts, such as the Prisoner's Dilemma [40, 41].

Cultural differences are one growing topic in apology studies because cultures have strong influences on the styles and meaning of apologies. In fact, a past study places Japan and the United States at opposite poles on well-established individualism–collectivism scales [42]. Because collectivist cultures foster a sense of shared agency, whereas individualist cultures prioritize personal accountability, such differences are

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

likely to alter how apologies are interpreted. Several studies focused on culture-specific apology behaviors [30, 43, 44], suggesting the need for appropriate apology strategies due to each cultural context.

Due to the advances of social robots that act in real settings, researchers in human-robot interaction have also focused on apology strategies for them. For example, researchers reported that the effectiveness of apology and compensation mitigates robot's failure [26]. Another study reported on the effectiveness of demonstrating self-awareness and ownership of robot's mistakes [27]. A recent study focused on appropriate bowing behaviors of a robot to express effective apologies [45]. However, in robots' apology context, there is less focus on cultural differences.

Therefore, in this study, we focused on how the effects of robot apologies are different between Japan and the United States (the U.S). The reasons why we selected the U.S are 1) as described above, Japan and the U.S. stand at opposite ends of well-established individualism–collectivism indices [42], which have influenced on the perceptions of robot apologies, and 2) the U.S. is the most extensively studied Western culture in social-robot research, then replicating Japanese findings with U.S. participants not only tests cross-cultural generalizability but also facilitates comparison with the bulk of existing HRI literature. Based on these considerations, we will discuss the detailed hypotheses and predictions in the next subsection.

2.2 Effects of multiple robots in various settings

Robotics researchers have been broadly investigating the effects of using multiple robots in various domains [46]: information-providing tasks [47-50], enhancing motivation [51-53], persuasive technology [54-56], product recommendations [3, 4, 57, 58], cute designs for robot behaviors [59], and robot-robot

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

interaction designs [60-64]. These studies can collectively be viewed as empirical tests of the Social Impact Theory [65] in different settings. The Social Impact Theory models social influence as a multiplicative function of the number, strength, and immediacy of influence sources, predicting that larger source sets exert a stronger impact. By manipulating only the "number" term, these studies showed the effectiveness of multi-robot settings. Following these findings, increasing the number of apologizing robots would amplify the effectiveness, too.

In fact, a previous study in Japan showed that apologies offered by multiple robots can be more effective than those delivered by a single robot [28]; perhaps, such multiple robot apologies would be perceived as more costly than a single robot apology. Although research outside Japan has not focused specifically on multi-robot apologies, other studies suggest that employing more than one robot is advantageous in various contexts: enhancing information-providing tasks [66], education support for children [50, 67-69], public exhibitions [49, 57], collaborative navigation [70], and managing peer pressure [56, 71-73]. Similarly, work within Japan has examined multiple-robot systems for such applications as drawing attention in advertising [3, 4, 57], boosting motivation and performance [52, 74], and addressing peer pressure [55]. Moreover, even though the only human-human interaction perspectives, a past study reported that numerosity effects are robust across cultural contexts [75]. Collectively, these studies imply that multi-robot interactions may be beneficial across different cultural settings.

Based on these considerations and the wider effects of the Social Impact Theory, we hypothesized that the enhanced acceptance of multi-robot apologies found in Japan will also manifest in other cultural contexts. We predicted that if two robots jointly apologized for a failure, U. S. participants would find these apologies more acceptable than those from a single robot, mirroring findings from Japan.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

Prediction 1: If multiple robots apologize for a failure/mistake, U.S. participants will accept the apologies more than just one robot's apology, similar to Japanese participants.

Prediction 2: If multiple robots apologize for a failure/mistake, U.S. participants will trust the apologies more than just one robot's apology, similar to Japanese participants.

2.3 Appropriate number of robots in apology

Following the Social Impact Theory [65], increasing the number of robots would simply increase their influence. Past studies reported that more robots strongly attracted peoples' attention [57], and six exhibited stronger peer pressure effects than two or four robots [55]; these studies suggested the effectiveness of increasing the number of robots.

However, we thought that increasing the number of robots is not always effective, and that there is an appropriate number for each task. For example, one study reported that too many robots may lower their acceptance of the robots in the context of the feeling of *kawaii* [59]. Another study argued that excessive apology has a risk of being perceived as an intentional action to get forgiveness [76]. From the perspective of Social Impact Theory [65], increasing numbers can sometimes weaken other elements, particularly immediacy (defined as the perceived closeness or intimacy of the communication). Following this context, an excessive number of robots or apologies may be inappropriate, as reported in the above studies; if the failure is relatively minor, an apology from too many robots may be perceived as insincere or regarded as an intentional action. Based on these considerations, we hypothesized that involving too

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

many robots when apologizing is unacceptable in both Japan and the U.S. We predicted that two robots would be received better than either four or six robots:

Prediction 3: Participants in Japan will forgive the robot's failure more when two robots apologize than when four or six apologize.

Prediction 4: Participants in the U.S. will forgive the robot's failure more when two robots apologize more than four or six apologize.

2.4 Summary of our study

This study investigated the above four hypotheses through two experiments. Experiment 1 tackled H1 to H2 by a test where participants perceived different acceptance of apology and trust toward apologies from a single/multiple robots, in both Japan and the U.S. For this purpose, we measured participants' feelings of forgiveness and trust by conducting a web-based survey. As described below, we employed data from a previous Japanese study [28] for comparison, and then we measured both items. Experiment 2 tackled H3 and H4 by a test where participants perceived different acceptance of apology toward apologies from two/four/six robots, in both Japan and the U.S. For this purpose, we simply measured participants' feelings of forgiveness by conducting an additional web-based survey, different from Experiment 1.

3. Experiment 1: Cultural differences of multiple robots' apologies between Japan and the U.S.

3.1 Visual stimuli and conditions

We employed the visual stimuli from a past study [28], merely translating the robot's spoken script into English. In each stimulus, a food-service robot (Pepper [77]) accidentally dropped a customer's order and offered an apology. Each video was approximately 30 seconds long, with 1286 × 762 pixel resolution at 30 fps. The

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

experiment employed a *number* factor; i.e., the number factor of robots that apologized (*one* vs. *two*). Moreover, we employed a *country* factor, i.e., *Japan* and *US*.

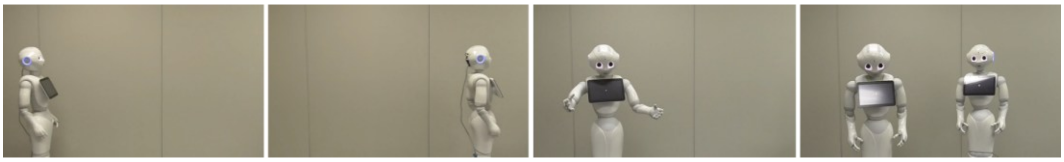
To illustrate the setup, two robots were stationed in a hypothetical restaurant environment. The secondary robot enters the frame from the left edge, moves across to the right edge, and then exits the frame (Fig.2 -a, b). The primary robot approached with an ice cream cone, saying, "Here you are." Unfortunately, it dropped the cone and apologized (Fig.2-c), "I'm so sorry," and offered a free replacement. We adopted the combined apology and compensation approach on the basis of prior work identifying suitable mitigation strategies for robot failures [26]. That study demonstrated that apology-only and compensation-only strategies vary in effectiveness depending on users' attitudes toward the robot, but such attitudes cannot be identified in advance in real-world settings. By pairing an apology with compensation, we aimed to accommodate both preference profiles and ensure broader applicability of the mitigating strategy. Another study also employed such a combination as a basic apology strategy [78], therefore, combining them would be rational. Here we described the differences between *one* and *two* conditions:

***One* condition:** Only the primary robot apologized; the secondary robot appeared briefly at the beginning, but not during the apology.

***Two* condition:** The secondary robot entered from the right and bowed in apology after the primary robot's failure, demonstrating a coordinated apology (Fig.2-d).

In both scenarios, the secondary robot did not speak to maintain consistent verbal apology content across the conditions.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563



(a) (b) (c) (d)

Figure 2 Screenshots from the visual stimuli in Experiment 1

3.2 Measurements

We measured the participants' forgiveness of the robot's failure (for prediction 1) and their perceived trust (for prediction 2) in the robot. For this purpose, we adapted several established scales: the forgiveness scale (the degree to forgiveness) [79], negative word-of-mouth (NWOM, the degree to want to spread negative word of mouth), performance trust from the Multi-Dimensional Measure of Trust (MDMT v2; reliable and competent subscales, trust from performance aspects) [80], moral trust from the MDMT v2 (ethical, transparent, and benevolent subscales, trust from moral aspects) [80], and intention to use (ITU, the degree of social acceptance) [81]. All items were rated on a 7-point scale (1 = most negative, 7 = most positive). We note that all English items were translated into Japanese and back-translated by a bilingual expert.

Because online surveys can introduce variability in participants' attention [82, 83], we incorporated two control questions to confirm that our respondents carefully watched the videos and followed instructions: 1) The video-content checks investigated whether the participants identified which item the robot dropped from among four options and what the robot did in the video among three options; those who answered incorrectly were excluded. 2) The instruction-compliance check determined whether participants left certain items blank following the instructions; those who disregarded this instruction were also excluded.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

3.3 Procedure

All the procedures were approved by the Advanced Telecommunication Research Review Boards (501-3). A consent form was presented on the initial webpage, and only participants who agreed to its terms were permitted to continue. Next participants were given an overview of the study, guidance for evaluating the videos, and a sound check to ensure that their audio playback was functioning adequately. Written descriptions provided context for imagining a cafeteria setting operated by multiple robots.

We employed a within-participant design. Participants viewed a video under an apology condition (one or two robots) and completed a relevant questionnaire. They then viewed the second video depicting the alternate condition and answered the same set of questions. The order of the two conditions was counterbalanced to minimize order effects. Participants completed the control questions following both videos to verify attention and instruction compliance.

3.4 Participants

Before conducting the experiment, we performed an a priori power analysis using G*power [84] (small effect size = 0.10, power = 0.80, and $\alpha = 0.05$). The required sample size was 200 participants. Therefore, first, we recruited 200 people in the United States through Amazon Mechanical Turk (AMT), anticipating potential exclusions. 77 identified as women, 121 as men, and 2 declined to provide gender information. After removing invalid responses, i.e., incomplete data or consistently identical ratings, 122 valid datasets remained (50 women, 71 men, and 1 who declined to specify). All participants received a nominal payment (under USD 5), regardless whether their responses were ultimately included. To examine cultural differences, we also employed the data from a previous Japanese study [28], which originally included 203 individuals

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

(101 women, 101 men, and 1 who declined to specify) and yielded 168 valid responses (81 women, 86 men, and 1 undisclosed). In total, the number of valid participants is 290, over the required sample size. Because the Japanese dataset had already been analyzed [28], our principal contribution in Experiment 1 is comparing those established results with responses from the new U.S. participants.

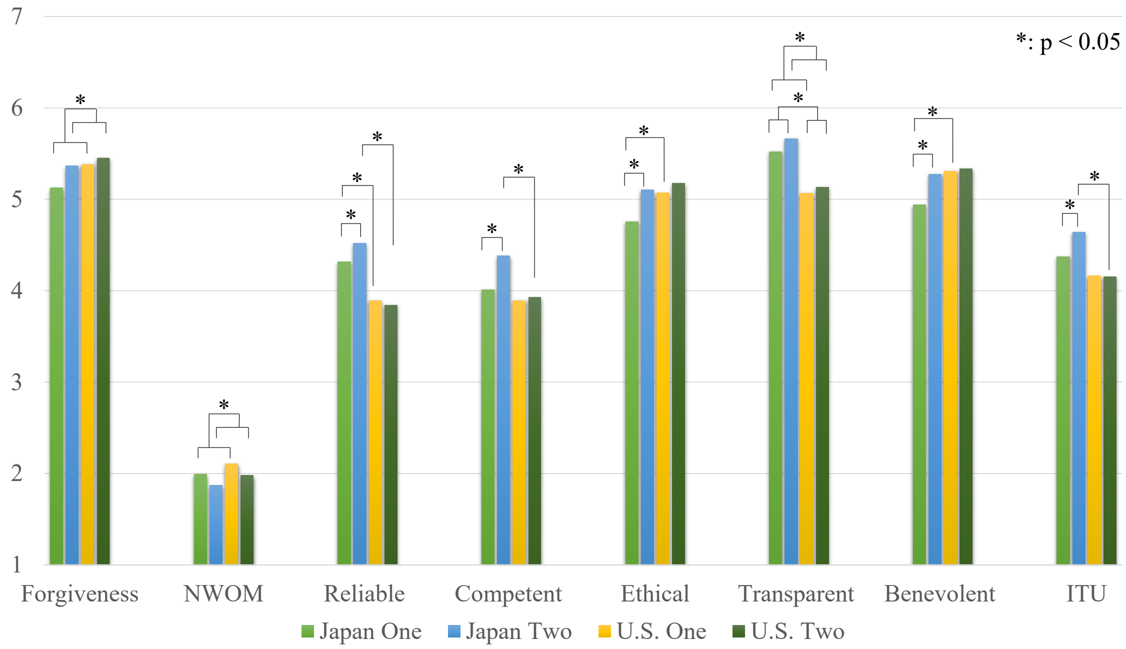


Figure 3. Questionnaire results from a previous study conducted in Japan [28] and additional data from the U.S. Where interaction effects reach significance, only the corresponding simple main effects are shown.

3.5 Results

Prior to the main analysis, Shapiro–Wilk tests were conducted to assess the normality of the distributions for each item. The results indicated that none of the distributions satisfied the assumption of normality ($p < .001$ for all conditions), thus we decided to use non-parametric methods for subsequent analyses. Therefore, we conducted an aligned rank transform (ART) ANOVA [85] with the *number* factor (*one* and *two*) and the *country* factor (*JP* and *US*) as factors for all questionnaire items (Fig.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

3). We employed the Holm method to correct the threshold of the p-values in the statistical testing. We note that although this study focused on predictions 1 and 2, which do not focus on the interaction effects between the factors of number and country, we conducted post hoc tests to explore the simple main effects when a significant interaction was found as an exploratory analysis for a deeper understanding.

For the forgiveness scale (Cronbach's alpha was 0.879), a significant main effect emerged for the *number* factor (*one* < *two,* $F(1, 288) = 12.808$, $p < 0.001$, partial $\eta^2 = 0.043$). No main effect was found for the *country* factor ($F(1, 288) = 1.451$, $p = 0.229$, partial $\eta^2 = 0.005$) or interaction between them ($F(1, 288) = 3.3234$, $p = 0.069$, partial $\eta^2 = 0.011$).

For the NWOM scale (a lower score is better, Cronbach's alpha was 0.903), a significant main effect emerged for the *number* factor (*two* < *one,* $F(1, 288) = 18.201$, $p < 0.001$, partial $\eta^2 = 0.059$). There was no main effect of the *country* factor ($F(1, 288) = 0.130$, $p = 0.719$, partial $\eta^2 = 0.001$) or significant interaction ($F(1, 288) = 0.018$, $p = 0.893$, partial $\eta^2 = 0.001$).

For the reliable subscale (Cronbach's alpha was 0.844), there were significant main effects for the *country* factor (*US* < *JP,* $F(1, 288) = 16.926$, $p < 0.001$, partial $\eta^2 = 0.056$) and interaction effects ($F(1, 288) = 9.523$, $p = 0.002$, partial $\eta^2 = 0.032$). We found no main effect of the *number* factor ($F(1, 288) = 1.537$, $p = 0.216$, partial $\eta^2 = 0.005$). Post hoc analyses using Wilcoxon rank sum tests to examine simple main effects of the *country* factor showed significant differences: *US's one* < *JP's one* ($p = 0.003$, $r = 0.175$), and *US's two* < *JP's two* ($p < 0.001$, $r = 0.274$). Post hoc analyses using Wilcoxon signed-rank sum test to examine simple main effects of the *number* factor showed a significant difference: *JP's one* < *JP's two* ($p < 0.001$, $r = 0.185$).

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

For the competent subscale (Cronbach's alpha was 0.890), there were significant main effects for *number* factor (*one* < *two,* $F(1, 288) = 18.128$, $p < 0.001$, partial $\eta^2 = 0.059$) and interaction effects ($F(1, 288) = 14.169$, $p < 0.001$, partial $\eta^2 = 0.047$). We found no main effect of the *country* factor ($F(1, 288) = 3.753$, $p = 0.054$, partial $\eta^2 = 0.013$). Post hoc analyses using Wilcoxon rank sum tests to examine simple main effects of the *country* factor showed a significant difference: *US's two* < *JP's two* ($p = 0.005$, $r = 0.167$). Post hoc analyses using Wilcoxon signed-rank sum test to examine simple main effects of the *number* factor showed a significant difference: *JP's one* < *JP's two* ($p < 0.001$, $r = 0.317$).

For the ethical subscale (Cronbach's alpha was 0.903), there were significant main effects for the *number* factor (*one* < *two,* $F(1, 288) = 38.651$, $p < 0.001$, partial $\eta^2 = 0.118$), the *country* factor (*US* < *JP,* $F(1, 288) = 4.708$, $p = 0.031$, partial $\eta^2 = 0.016$), and interaction effects ($F(1, 288) = 11.250$, $p < 0.001$, partial $\eta^2 = 0.038$). Post hoc analyses using Wilcoxon rank sum tests to examine simple main effects of the *country* factor showed a significant difference: *JP's one* < *US's one* ($p = 0.005$, $r = 0.164$). Post hoc analyses using Wilcoxon signed-rank sum test to examine simple main effects of the *number* factor showed a significant difference: *JP's one* < *JP's two* ($p < 0.001$, $r = 0.344$).

For the transparent subscale (Cronbach's alpha was 0.866), there were significant main effects for the *number* factor (*one* < *two,* $F(1, 288) = 12.673$, $p < 0.001$, partial $\eta^2 = 0.042$) and the *country* factor (*JP* < *US,* $F(1, 288) = 15.266$, $p = 0.001$, partial $\eta^2 = 0.050$). The interaction effect was not statistically significant ($F(1, 288) = 0.866$, $p = 0.353$, partial $\eta^2 = 0.003$).

For the benevolent subscale (Cronbach's alpha was 0.930), there were significant main effects for the *number* factor (*one* < *two,* $F(1, 288) = 25.953$, $p < 0.001$,

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

partial $\eta^2 = 0.083$), the *country* factor (*US* < *JP*, $F(1, 288) = 4.866$, $p = 0.028$, partial $\eta^2 = 0.017$) and interaction effects ($F(1, 288) = 19.137$, $p < 0.001$, partial $\eta^2 = 0.062$). Post hoc analyses using Wilcoxon rank sum tests to examine simple main effects of the *country* factor showed a significant difference: *JP's one* < *US's one* ($p = 0.001$, $r = 0.191$). Post hoc analyses using Wilcoxon signed-rank sum test to examine simple main effects of the *number* factor showed a significant difference: *JP's one* < *JP's two* ($p < 0.001$, $r = 0.348$).

For the ITU subscale (Cronbach's alpha was 0.937), there were significant main effects for *number* factor (*one* < *two*, $F(1, 288) = 17.062$, $p < 0.001$, partial $\eta^2 = 0.056$) and interaction effects ($F(1, 288) = 8.849$, $p = 0.003$, partial $\eta^2 = 0.030$). There was no main effect for the *country* factor ($F(1, 288) = 3.288$, $p = 0.071$, partial $\eta^2 = 0.011$). Post hoc analyses using Wilcoxon rank sum tests to examine simple main effects of the *country* factor showed a significant difference: *US's two* < *JP's two* ($p = 0.013$, $r = 0.145$). Post hoc analyses using Wilcoxon signed-rank sum test to examine simple main effects of the *number* factor showed a significant difference: *JP's one* < *JP's two* ($p < 0.001$, $r = 0.231$).

3.6 Discussion of Experiment 1

In both Japan and the United States, participants generally responded more positively to apologies offered by multiple robots than those delivered by a single robot, as evidenced by higher scores on the forgiveness and NWOM measures; thus, prediction 1 was supported.

However, the exploratory analysis about the interaction effects showed that the effect of the *number* factor on the trust subscales varied between the two countries. Compared to Japanese participants, Americans showed smaller differences in trust-related ratings between the single- and multi-robot apology conditions, where the only transparent subscale was an exception; thus, prediction 2 was not supported.

This study explored how participants in two different countries respond to apologies issued by multiple robots. Although cultural differences in the perceptions of apologies have been reported [29-32], our findings indicate that, overall, multi-robot apologies are more favorably received than single-robot apologies in both cultural contexts. These outcomes offer an additional perspective on the potential advantages of using multiple robots, specifically for managing mistakes and preserving positive user perceptions. They also provide additional evidence of the validity of the Social Impact Theory [65] in the context of robot apology.

Interestingly, differences emerged in the participants' trust-related evaluations. Americans appeared less influenced by having multiple robots apologize than their Japanese counterparts, potentially reflecting distinctions between a culture emphasizing individual agency (the United States) and another emphasizing collective agency (Japan) [42]. In the former, an apology delivered by an additional robot may have a limited impact on how a participant evaluates the robot that made a mistake. By contrast, in the latter, people may view robots as members of an interconnected group, leading to a more pronounced effect when another "in-group" robot contributes to the apology. Indeed, prior research suggests that Japanese people are more likely to apologize for failures in which they were not directly involved [30]. Thus, one alternative explanation is that Japanese participants viewed the absence of a joint apology (i.e., when the secondary robot did not apologize) less favorably, thus reducing certain trust ratings.

As an additional analysis, we investigated the correlations between all measures; the results showed that all measures showed significant correlations with each other.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

Although this analysis could not provide clear causal relationships between them, we thought that apologies involving multiple robots increased perceived trust and forgiveness, as well as decreased NWOM. Investigating the causal relationship between the number of robots and the effects of apologies would be an interesting research topic. We note that to avoid redundant explanations about them here, we have prepared the full correlation tables, as well as the full statistical results of the ART ANOVA, in the supplementary data file.

4 Experiment 2: Effects of the number of robots on apologies between Japan and the U.S.

4.1 Visual stimuli and conditions

We modified the visual stimuli of Experiment 1, adding more robots for the apology timing. We employed two, four, and six robots because such numbers were also employed in a past study that investigated the effects of the number of robots [55]. The video starts with the primary robot bringing an ice cream cone; the same action is used to apologize after it is dropped. After that, a different number of robots appears, depending on the conditions, and they bow without any verbal information. We prepared six videos (the *number* factor: two, four, and six robots, the *country* factor: *Japan* and *U.S.*). Each video was approximately 25 seconds long, with 1920 × 780 pixel resolution at 30 fps.

4.2 Measurements

We measured the participants' forgiveness of the robot's failure. We again adapted established scales: the forgiveness scale [79] and the negative word-of-mouth (NWOM). All items were rated on a 7-point scale (1 = most negative, 7 = most

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

positive). We employed two control questions to confirm that respondents watched the videos carefully and followed instructions.

4.3 Procedure

The experiment procedure of Experiment 2 was identical as Experiment 1. All the procedures were approved by the Advanced Telecommunication Research Review Boards (501-3). A consent form was presented on the initial webpage, and only participants who agreed to its terms were permitted to continue. We again employed a within-participant design for the *number* factor. Thus, participants viewed a video under one apology condition (*two*, *four,* or *six*) and completed a relevant questionnaire. They then viewed the rest of the videos and answered the same set of questions. The order of the three conditions was counterbalanced to minimize order effects. Participants completed the control questions following both videos to verify attention and instruction compliance.

4.4 Participants

Before conducting the experiment, we performed an a priori power analysis using G*power [84] (small effect size = 0.10, power = 0.80, and $\alpha = 0.05$). The required sample size was 164 participants. Therefore, first, we recruited 344 Japanese people through a local company. After removing invalid responses, i.e., incomplete data or consistently identical ratings, 265 valid datasets remained (134 women, 131 men). Also, 300 individuals in the United States were recruited through Amazon Mechanical Turk (AMT). After removing invalid responses, i.e., incomplete data or consistently identical ratings, 273 valid datasets remained (150 women, 122 men, and 1 who declined to specify). In total, the number of valid participants is 538, over the required sample size.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

All participants received a nominal payment (under USD 5), regardless whether their responses were ultimately included.

4.5 Results

Prior to the main analysis, Shapiro–Wilk tests were conducted to assess the normality of the distributions for each item. The results indicated that none of the distributions satisfied the assumption of normality ($p < .001$ for all conditions), thus we decided to use non-parametric methods for subsequent analyses. Therefore, we conducted an aligned rank transform (ART) ANOVA [85] with the *number* of robots delivering an apology (*two, four* or *six*) and the *country* factor (*JP* and *US*) as factors for all questionnaire items (Fig. 4). We employed the Holm method to correct the threshold of the p-values in the statistical testing. Similar to Experiment 1, we conducted the analysis of the interaction effects as an exploratory analysis for a deeper understanding.

For the forgiveness scale (Cronbach's alpha was 0.937), significant main effects emerged for the *number* factor ($F(2, 1072) = 106.792$, $p < .001$, partial $\eta^2 = 0.166$), the *country* factor (*US < JP,* $F(1, 536) = 23.828$, $p < 0.001$, partial $\eta^2 = 0.043$) and significant interaction effect ($F(2, 1072) = 51.034$, $p < 0.001$, partial $\eta^2 = 0.087$). Post hoc analyses using Wilcoxon rank sum tests to examine simple main effects of the *country* factor showed significant differences: *JP's four < US's four* ($p < 0.001$, $r = 0.214$), and *JP's six < US's six* ($p < 0.001$, $r = 0.284$). Post hoc analyses using Wilcoxon signed-rank sum test to examine simple main effects of the *number* factor showed a significant difference: *JP's four < JP's two* ($p < 0.001$, $r = 0.357$), *JP's six < JP's two* ($p < 0.001$, $r = 1.000$), *JP's six < JP's four* ($p < 0.001$, $r = 1.000$), *US's six < US's two* ($p < 0.001$, $r = 0.192$), and *US's six < US's four* ($p < 0.001$, $r = 0.175$).

For the NWOM scale (a lower score is better, Cronbach's alpha was 0.917), significant main effects emerged for the *number* factor ($F(2, 1072) = 46.324$, $p < 0.001$,

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

partial $\eta^2 = 0.080$) and significant interaction effect ($F(2, 1072) = 17.616$, $p < 0.001$, partial $\eta^2 = 0.032$). No significant main effect was found for the *country* factor ($F(1, 536) = 0.0326$, $p = 0.857$, partial $\eta^2 < 0.001$). Post hoc analyses using Wilcoxon rank sum tests to examine simple main effects of the *country* factor showed a significant difference: *JP's two* < *US's two* ($p < 0.001$, $r = 0.143$). Post hoc analyses using Wilcoxon signed-rank sum test to examine simple main effects of the *number* factor showed a significant difference: *JP's two* < *JP's four* ($p < 0.001$, $r = 0.309$), *JP's two* < *JP's six* ($p < 0.001$, $r = 1.000$), *JP's four* < *JP's six* ($p < 0.001$, $r = 0.228$), *US's two* < *US's four* ($p = 0.014$, $r = 0.105$), *US's two* < *US's six* ($p < 0.001$, $r = 0.195$), and *US's four* < *US's six* ($p = 0.026$, $r = 0.095$).

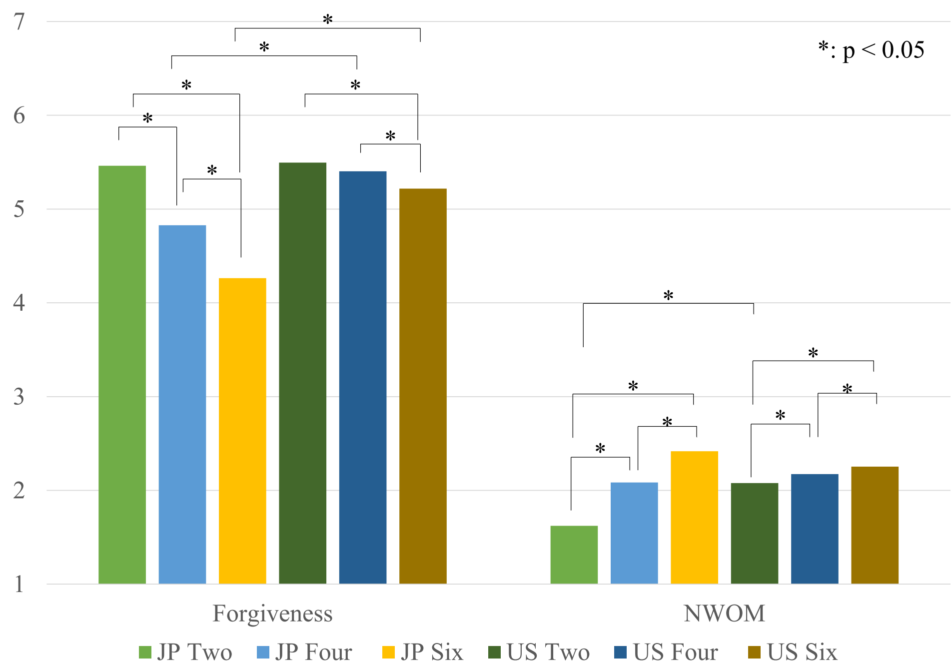


Figure 4. Questionnaire results of Experiment 2. Where interaction effects reach significance, only the corresponding main effects are shown.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

4.6 Discussion of Experiment 2

Next, we summarize the experiment results. In Japan, the participants generally responded more positively to apologies offered by two robots than those delivered by four and six, as evidenced by higher scores on forgiveness and lower scores on the NWOM measures. On the other hand, in the U.S., participants responded more positively to apologies offered by two robots than those delivered by six robots but not four. Therefore, prediction 3 was supported; prediction 4 was partially supported.

These results showed that increasing the number of apologizing robots tended to worsen evaluations of the apology in both countries. Notably, the magnitude of this deterioration was larger among Japanese participants than among U.S. participants. A straightforward conformity-based account does not satisfactorily explain this pattern. Popular stereotypes sometimes portray Japanese people and culture as especially conformist compared with the U.S. On that view, larger apologizing groups should have been relatively more persuasive in Japan. However, past research has reported broadly similar levels of conformity in Japan and the U.S. [86]. Accordingly, a conformity explanation appears insufficient for the present findings.

There are two possible explanations for this. The first is negative attitudes toward excessive apologies. A study conducted in Japan showed that excessive apologies by a robot for minor mistakes (e.g., bowing longer or deeper for a simple error) receive lower evaluations [45]. Although this past study focused only on bow depth and duration, in our setting, increasing the number of robots may similarly have made the apologies seem excessive, leading to lower evaluations. Another possibility is that Japanese participants perceived the multi-robot apology as substantively insufficient. In our experiment, only the primary robot verbalized the apology, whereas the additional robots merely bowed after a delay. As the group grew, the proportion of

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

light apologies increased, possibly leading participants to judge the group's apology as insufficient apologies overall.

To distinguish between these possibilities, future work should include questionnaire items that assess perceived excessiveness and sufficiency of the apology, and/or add conditions that standardize the content and timing of multiple robots' apologies (e.g., all robots speak or all bow synchronously). These extensions would be a valuable direction for future research and would deepen the present findings.

We additionally investigated the correlations between all measures; the results showed that all measures showed significant correlations with each other similar to Experiment 1. In this experiment, we did not measure perceived trust, but a similar phenomenon might have appeared. To avoid redundant explanations about them here, we prepared the full correlation tables as well as the full statistical results of ART ANOVA in the supplementary data file.

5 General Discussion

5.1 Implication

Through two experiments, our research consistently demonstrated that two-robot apologies showed better performance in the context of a failure in a restaurant setting, which is typically chosen to investigate the effects of robot apology strategy [26]. A past study reported that in various countries, costly apologies are generally more effective than no-cost apologies [87], although our study suggests that simply increasing the cost of apologies is ineffective, at least in robot apology settings.

Related to this point of view, note that the effectiveness of increasing the number of robots is limited in the context of apologizing, unlike contexts of peer pressure [55] or advertising [57]. More robots increase the social pressures or

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

attractiveness and strengthen these social influences. Of course, such strong social pressures are not appropriate for apology settings. A sincere attitude is necessary for an apology, and we must avoid designing behavior that implies any sense of coercion to the acceptance of apologies. In this study, we focused on how the number of robots affected the cost of apologizing; interesting future work might examine the impact of verbal and non-verbal information on such costs and the perceived coercion of apologies.

5.2 Limitations

We note several limitations of our study. First, it only examined one type of robot (Pepper), specific numbers of robots (two, four, or six), and a specific failure (dropping an ice cream cone). Different results might emerge if more humanlike androids [88-90] or other robotic platforms were used [91-93], as well as more robots in the apologies. Second, although we incorporated data from both Japan and the United States, additional research in other cultural settings could generate broader insights into how social and cultural norms shape perceptions of multi-robot apologies. Third, we did not use the “not applicable” option in the use of the Multi-Dimensional Measure of Trust (MDMT) v2 questionnaires, which would limit the validity of a part of the experimental results. Such a broader approach might facilitate the design of robots that are better aligned with the diverse values and expectations of users around the globe. Fourth, we did not measure participants' mind-perception of the robot, even though such perceptions may influence the acceptance of robot apologies. Because we used a within-participant design, i.e., every participant experienced each condition, any stable tendency to attribute more or less “mind” to the robot was held constant. Nevertheless, examining how individual differences in mind-perception moderate responses to robot apologies remains an important direction for future research. Finally, our findings do

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

not reflect a pure apology effect because we employed the combination approach of apology and compensation. Nevertheless, in commercial service contexts, apologies are seldom offered on their own; they are typically accompanied by a concrete form of redress (e.g., a voucher, refund, or replacement). Therefore, we thought that our manipulation would provide knowledge about a more realistic situation than a stand-alone apology.

6 Conclusion

This study investigated the cross-cultural perceptions of multi-robot apologies in Japan and the United States. The first experiment compared one- and multiple-robot apologies, revealing that participants in both countries generally found apologies from multiple robots more acceptable than those from a single robot, particularly in terms of forgiveness and negative word-of-mouth (NWOM). However, trust-related measures showed that U.S. participants were somewhat less influenced by the number of apologizing robots than their Japanese counterparts.

To extend these findings, our second experiment further explored varying numbers of robots (two, four, and six) in both countries. Across this expanded range, two-robot apologies consistently elicited the most favorable responses, suggesting that simply increasing the number of apologizing robots beyond two does not yield additional benefits. Taken together, these results underscore the importance of carefully calibrating the number of robots who are apologizing, while also accounting for cultural differences in how trust is formed. This knowledge can inform the design of social robots that effectively restore user satisfaction and confidence following failures.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

Acknowledgement

This work was partially supported by JST Moonshot R&D Grant Number JPMJMS2011 (Evaluation), JST CREST Grant, Number JPMJCR18A1, Japan (Experiment 1), and JSPS KAKENHI Grant Numbers JP22H03895 (Experiment 2), JP23K16980 (Preparation materials) and JP24K21327 (Writing). We thank ChatGPT (OpenAI, GPT-4o) for the initial English language editing and proofreading of the manuscript. The final version of the manuscript was reviewed and further edited by a professional editing service.

Disclosure Statement

No potential conflict of interest was reported by the author(s). The funders had no role in study design, data collection and analysis, decision to publish, or manuscript preparation.

Data Availability Statement

Data are available on request from the corresponding author.

References

[1] Y. Jiang, and J. Wen, “Effects of COVID-19 on hotel marketing and management: a perspective article,” *International Journal of Contemporary Hospitality Management*, 2020.

[2] S. S. Kim, J. Kim, F. Badu-Baiden, M. Giroux, and Y. Choi, “Preference for robot service or human service in hotels? Impacts of the COVID-19 pandemic,” *International Journal of Hospitality Management,* vol. 93, pp. 102795, 2021.

[3] S. Song, J. Baba, Y. Okafuji, J. Nakanishi, Y. Yoshikawa, and H. Ishiguro, “Out for In! Empirical Study on the Combination Power of Two Service Robots for Product Recommendation,” in Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction, Stockholm, Sweden, pp. 408–416, 2023.

[4] T. Iwamoto, J. Baba, J. Nakanishi, K. Hyodo, Y. Yoshikawa, and H. Ishiguro, “Playful Recommendation: Sales Promotion That Robots Stimulate Pleasant Feelings Instead of Product Explanation,” *IEEE Robotics and Automation Letters,* vol. 7, no. 4, pp. 11815-11822, 2022.

[5] M. Shiomi, K. Shinozawa, Y. Nakagawa, T. Miyashita, T. Sakamoto, T. Terakubo, H. Ishiguro, and N. Hagita, “Recommendation Effects of a Social Robot for Advertisement-Use Context in a Shopping Mall,” *International Journal of Social Robotics,* vol. 5, no. 2, pp. 251-262, 2013.

[6] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita, “A communication robot in a shopping mall,” *Robotics, IEEE Transactions on,* vol. 26, no. 5, pp. 897-913, 2010.

[7] P. Kong, H. Cornet, and F. Frenkler, "Personas and emotional design for public service robots: A case study with autonomous vehicles in public transportation," in 2018 international conference on cyberworlds (cw), pp. 284-287, 2018.

[8] M. Tonkin, J. Vitale, S. Herse, M.-A. Williams, W. Judge, and X. Wang, "Design methodology for the ux of hri: A field study of a commercial social robot at an airport," in Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, pp. 407-415, 2018.

[9] M. Shiomi, D. Sakamoto, T. Kanda, C. T. Ishi, H. Ishiguro, and N. Hagita, "Field Trial of a Networked Robot at a Train Station," *International Journal of Social Robotics,* vol. 3, no. 1, pp. 27-40, 2010.

[10] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, "Interactive Humanoid Robots for a Science Museum," *IEEE Intelligent Systems*, no. 2, pp. 25-32, 2007.

[11] P. K. Nikolić, and M. R. M. Tomari, "Robot-Robot Interaction, Toward New Conversational Artificial Intelligence Aesthetic," in 10th International Conference on Digital and Interactive Arts, Aveiro, Portugal, Portugal, pp. Article 15, 2022.

[12] A. Nijholt, "Robotic Stand-Up Comedy: State-of-the-Art," in Distributed, Ambient and Pervasive Interactions: Understanding Humans, Cham, pp. 391-410, 2018.

[13] T. Iio, Y. Yoshikawa, M. Chiba, T. Asami, Y. Isoda, and H. Ishiguro, "Twin-Robot Dialogue System with Robustness against Speech Recognition Failure in Human-Robot Dialogue with Elderly People," *Applied Sciences,* vol. 10, no. 4, pp. 1522, 2020.

[14] H. Sumioka, N. Yamato, M. Shiomi, and H. Ishiguro, "A minimal design of a human infant presence: a case study toward interactive doll therapy for older adults with dementia," *Frontiers in Robotics and AI,* vol. 8, pp. 633378, 2021.

[15] N. Yamato, H. Sumioka, H. Ishiguro, M. Shiomi, and Y. Kohda, "Technology Acceptance Models from Different Viewpoints of Caregiver, Receiver, and Care Facility Administrator: Lessons from Long-Term Implementation Using Baby-Like Interactive Robot for Nursing Home Residents with Dementia," *Journal of Technology in Human Services*, pp. 1-26, 2023.

[16] M. Natarajan, and M. Gombolay, "Effects of anthropomorphism and accountability on trust in human robot interaction," in Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, pp. 33-42, 2020.

[17] T. Law, and M. Scheutz, "Trust: Recent concepts and evaluations in human-robot interaction," *Trust in human-robot interaction*, pp. 27-57, 2021.

[18] 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.

[19] K. Okuoka, K. Enami, M. Kimoto, and M. Imai, "Multi-device trust transfer: Can trust be transferred among multiple devices?," *Frontiers in Psychology,* vol. 13, pp. 920844, 2022.

[20] F. Fossa, and I. Sucameli, "Gender bias and conversational agents: an ethical perspective on social robotics," *Science and Engineering Ethics,* vol. 28, no. 3, pp. 23, 2022.

[21] E. K. Duarte, M. Shiomi, A. Vinel, and M. Cooney, "Trust in Robot Self-Defense: People Would Prefer a Competent, Tele-Operated Robot That Tries to Help," in 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 2447-2453, 2023.

[22] S. Anjomshoae, A. Najjar, D. Calvaresi, and K. Främling, "Explainable agents and robots: Results from a systematic literature review," in 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019, pp. 1078-1088, 2019.

[23] T. Nomura, "Robots and gender," *Gender and the Genome,* vol. 1, no. 1, pp. 18-25, 2017.

[24] R. Wullenkord, and F. Eyssel, "Societal and ethical issues in HRI," *Current Robotics Reports,* vol. 1, pp. 85-96, 2020.

[25] M. Maroto-Gómez, F. Alonso-Martín, M. Malfaz, Á. Castro-González, J. C. Castillo, and M. Á. Salichs, "A Systematic Literature Review of Decision-Making and Control Systems for Autonomous and Social Robots," *International Journal of Social Robotics,* vol. 15, no. 5, pp. 745-789, 2023.

[26] M. K. Lee, S. Kiesler, J. Forlizzi, S. Srinivasa, and P. Rybski, "Gracefully mitigating breakdowns in robotic services," in Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on, pp. 203-210, 2010.

[27] D. Cameron, S. de Saille, E. C. Collins, J. M. Aitken, H. Cheung, A. Chua, E. J. Loh, and J. Law, "The effect of social-cognitive recovery strategies on likability, capability and trust in social robots," *Computers in Human Behavior,* vol. 114, pp. 106561, 2021.

[28] Y. Okada, M. Kimoto, T. Iio, K. Shimohara, and M. Shiomi, "Two is better than one: Apologies from two robots are preferred," *Plos one,* vol. 18, no. 2, pp. e0281604, 2023.

[29] N. Sugimoto, "A Japan-US comparison of apology styles," *Communication Research,* vol. 24, no. 4, pp. 349-369, 1997.

[30] W. W. Maddux, P. H. Kim, T. Okumura, and J. M. Brett, "Cultural differences in the function and meaning of apologies," *International negotiation,* vol. 16, no. 3, pp. 405-425, 2011.

[31] R. Gee, and Y. Murata, "Comparing Malaysian apologies and responses to accusations with those of Japanese, British and Canadians: what is Malaysian pragmatic style?," *Asian Englishes,* vol. 16, no. 1, pp. 19-31, 2014.

[32] N. Isagozawa, and K. Fuji, "An Exploratory Study of Cultural Differences in Behavior and Intentions in Workplace Apologies: With the Perspective of Responsibility," *Japanese Psychological Research*, 2023.

[33] M. Shiomi, T. Hirayama, M. Kimoto, T. Iio, and K. Shimohara, "Two is Better than One: Cultural Differences in the Number of Apologizing Robots in the U.S. and Japan," in 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN), pp. 351-356, 2024.

[34] H. Wagatsuma, and A. Rosett, "The implications of apology: Law and culture in Japan and the United States," *Law & Soc'y Rev.,* vol. 20, pp. 461, 1986.

[35] J. Abeler, J. Calaki, K. Andree, and C. Basek, "The power of apology," *Economics Letters,* vol. 107, no. 2, pp. 233-235, 2010.

[36] L. A. Martinez-Vaquero, T. A. Han, L. M. Pereira, and T. Lenaerts, "Apology and forgiveness evolve to resolve failures in cooperative agreements," *Scientific reports,* vol. 5, no. 1, pp. 1-12, 2015.

[37] Y. Ohtsubo, M. Matsunaga, T. Himichi, K. Suzuki, E. Shibata, R. Hori, T. Umemura, and H. Ohira, "Costly group apology communicates a group's sincere "intention"," *Social Neuroscience,* vol. 15, no. 2, pp. 244-254, 2020.

[38] P. H. Kim, K. T. Dirks, C. D. Cooper, and D. L. Ferrin, "When more blame is better than less: The implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based trust violation,"

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 23 Oct 2025, available at: https://doi.org/10.1080/01691864.2025.2571563

Organizational Behavior and Human Decision Processes, vol. 99, no. 1, pp. 49-65, 2006.
[39] K. Sharma, F. D. Schoorman, and G. A. Ballinger, "How Can It Be Made Right Again? A Review of Trust Repair Research," *Journal of Management,* vol. 49, no. 1, pp. 363-399, 2023.
[40] T. A. Han, L. M. Pereira, F. C. Santos, and T. Lenaerts, "Why is it so hard to say sorry? evolution of apology with commitments in the iterated Prisoner's Dilemma," in Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, pp. 177-183, 2013.
[41] T. Han, L. Pereira, F. Santos, and T. Lenaerts, "Why is it so hard to say sorry," in BNAIC 2013: Proceedings of the 25th Benelux Conference on Artificial Intelligence, Delft, The Netherlands, November 7-8, pp. 4-9, 2013.
[42] G. Hofstede, "Culture's consequences: Comparing values, behaviors, institutions and organizations across nations," *International Educational and Professional*, 2001.
[43] M. Amri, "Ojigi: The Ethics of Japanese Community's Nonverbal Language," in Social Sciences, Humanities and Education Conference (SoSHEC 2019), pp. 305-308, 2019.
[44] M. Burdelski, ""I'm sorry, flower": Socializing apology, relationships, and empathy in Japan," *Pragmatics and Society,* vol. 4, no. 1, pp. 54-81, 2013.
[45] M. Shiomi, T. Hirayama, M. Kimoto, T. Iio, and K. Shimohara, "Modeling of Bowing Behaviors in Apology Based on Severity," *IEEE Robotics and Automation Letters,* vol. 9, no. 11, pp. 10169-10176, 2024.
[46] M. Shiomi, "A systematic survey of multiple social robots as a passive- and interactive-social medium," *Advanced Robotics,* vol. 38, no. 7, pp. 1-15, 2023.
[47] H. Itahara, M. Kimoto, T. Iio, K. Shimohara, and M. Shiomi, "How Does Exposure to Changing Opinions or Reaffirmation Opinions Influence the Thoughts of Observers and Their Trust in Robot Discussions?," *Applied Sciences,* vol. 13, no. 1, pp. 585, 2023.
[48] J. Nakanishi, I. Kuramoto, J. Baba, K. Ogawa, Y. Yoshikawa, and H. Ishiguro, "Continuous Hospitality with Social Robots at a hotel," *SN Applied Sciences,* vol. 2, no. 3, pp. 452, 2020.
[49] M. Aizawa, and H. Umemuro, "Behavioral Design of Guiding Agents to Encourage their Use by Visitors in Public Spaces," in Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, Boulder, CO, USA, pp. 247–251, 2021.
[50] A.-M. Velentza, N. Fachantidis, and S. Pliasa, "Which One? Choosing Favorite Robot After Different Styles of Storytelling and Robots' Conversation," *Frontiers in Robotics and AI,* vol. 8, 2021.
[51] M. Alemi, A. Taheri, A. Shariati, and A. Meghdari, "Social Robotics, Education, and Religion in the Islamic World: An Iranian Perspective," *Science and Engineering Ethics,* vol. 26, no. 5, pp. 2709-2734, 2020.
[52] M. Shiomi, S. Okumura, M. Kimoto, T. Iio, and K. Shimohara, "Two is better than one: Social rewards from two agents enhance offline improvements in motor skills more than single agent," *PloS one,* vol. 15, no. 11, pp. e0240622, 2020.
[53] M. Shiomi, Y. Tamura, M. Kimoto, T. Iio, R. Akahane-Yamada, and K. Shimohara, "Two is better than one: verification of the effect of praise from two robots on pre-school children's learning time," *Advanced Robotics,* vol. 35, no. 19, pp. 1132-1144, 2021.

[54] M. Shiomi, and N. Hagita, "Do Synchronized Multiple Robots Exert Peer Pressure?," in Proceedings of the Fourth International Conference on Human Agent Interaction, Biopolis, Singapore, pp. 27-33, 2016.

[55] M. Shiomi, and N. Hagita, "Do the number of robots and the participant's gender influence conformity effect from multiple robots?," *Advanced Robotics,* vol. 33, no. 15-16, pp. 756-763, 2019.

[56] M. Hashemian, A. Paiva, S. Mascarenhas, P. A. Santos, and R. Prada, "The power to persuade: a study of social power in human-robot interaction," in 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 1-8, 2019.

[57] J. Amada, Y. Okafuji, K. Matsumura, J. Baba, and J. Nakanishi, "Investigating the crowd-drawing effect, on passersby, of pseudo-crowds using multiple robots," *Advanced Robotics,* vol. 37, no. 6, pp. 423-432, 2023.

[58] Y. Kimura, E. Anzai, N. Saiwaki, and M. Shiomi, "Two is better than one: descriptions by multiple robots strengthen the feeling of kawaii toward objects," in 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN), pp. 485-490, 2024.

[59] M. Shiomi, R. Hayashi, and H. Nittono, "Is two cuter than one? number and relationship effects on the feeling of kawaii toward social robots," *PLOS ONE,* vol. 18, no. 10, pp. e0290433, 2023.

[60] Y. Okada, R. Taniguchi, A. Tatsumi, M. Okubo, M. Kimoto, T. Iio, K. Shimohara, and M. Shiomi, "Effects of Touch Behaviors and Whispering Voices in Robot-Robot Interaction for Information Providing Tasks," in 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 7-13, 2020.

[61] T. Hirayama, Y. Okada, M. Kimoto, T. Iio, K. Shimohara, and M. Shiomi, "Speed effects in touching behaviours: impact on perceived relationships in robot-robot interactions," *Advanced Robotics*, pp. 1-18.

[62] M. Shiomi, M. Kakio, and T. Miyashita, "Who Should Speak? Voice Cue Design for a Mobile Robot Riding in a Smart Elevator," in 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN), pp. 2023-2028, 2024.

[63] N. Karatas, S. Yoshikawa, P. R. S. De Silva, and M. Okada, "How Multi-Party Conversation Can Become an Effective Interface While Driving," *The Transactions of Human Interface Society,* vol. 20, no. 3, pp. 371-388, 2018.

[64] N. Karatas, S. Tamura, M. Fushiki, and M. Okada, "Multi-party Conversation of Driving Agents: The Effects of Overhearing Information on Lifelikeness and Distraction," in Proceedings of the 6th International Conference on Human-Agent Interaction, Southampton, United Kingdom, pp. 84–91, 2018.

[65] B. Latané, "The psychology of social impact," *American Psychologist,* vol. 36, no. 4, pp. 343-356, 1981.

[66] A. M. Velentza, D. Heinke, and J. Wyatt, "Human Interaction and Improving Knowledge through Collaborative Tour Guide Robots," in 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 1-7, 2019.

[67] P. Wicke, and T. Veale, "The Show Must Go On: On the Use of Embodiment, Space and Gesture in Computational Storytelling," *New Generation Computing,* vol. 38, no. 4, pp. 565-592, 2020.

[68] P. Soleiman, H. Moradi, B. Mehralizadeh, H. Ameri, R. I. Arriaga, H. R. Pouretemad, N. Baghbanzadeh, and L. K. Vahid, "Fully robotic social

environment for teaching and practicing affective interaction: Case of teaching emotion recognition skills to children with autism spectrum disorder, a pilot study," *Frontiers in Robotics and AI,* vol. 10, 2023.

[69] A. Esfandbod, Z. Rokhi, A. F. Meghdari, A. Taheri, Z. Soleymani, M. Alemi, and M. Karimi, "Fast mapping in word-learning: A case study on the humanoid social robots' impacts on Children's performance," *International Journal of Child-Computer Interaction,* vol. 38, pp. 100614, 2023.

[70] X. Z. Tan, S. Reig, E. J. Carter, and A. Steinfeld, "From one to another: how robot-robot interaction affects users' perceptions following a transition between robots," in Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction, Daegu, Republic of Korea, pp. 114–122, 2020.

[71] J. Brandstetter, P. Racz, C. Beckner, E. B. Sandoval, J. Hay, and C. Bartneck, "A peer pressure experiment: Recreation of the Asch conformity experiment with robots," in Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on, pp. 1335-1340, 2014.

[72] N. Salomons, M. v. d. Linden, S. S. Sebo, and B. Scassellati, "Humans Conform to Robots: Disambiguating Trust, Truth, and Conformity," in Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA, pp. 187–195, 2018.

[73] A.-L. Vollmer, R. Read, D. Trippas, and T. Belpaeme, "Children conform, adults resist: A robot group induced peer pressure on normative social conformity," *Science Robotics,* vol. 3, no. 21, pp. eaat7111, 2018.

[74] Y. Tamura, M. Shiomi, M. Kimoto, T. Iio, K. Shimohara, and N. Hagita, "Robots as an interactive-social medium in storytelling to multiple children," *Interaction Studies,* vol. 22, no. 1, pp. 110-140, 2021.

[75] R. Bond, and P. B. Smith, "Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task," *Psychological bulletin,* vol. 119, no. 1, pp. 111, 1996.

[76] J. C. Lee, and H. Chang, "Effect of Elaborateness of Apology on Subsequent Disciplinary Action Considering Outcome Severity and Favorable Reputation as Moderators," *Business Communication Research and Practice,* vol. 4, no. 2, pp. 104-114, 2021.

[77] A. K. Pandey, and R. Gelin, "A Mass-Produced Sociable Humanoid robot. Pepper: The first machine of its kind," *IEEE Robotics & Automation Magazine,* vol. 25, no. 3, pp. 40-48, 2018.

[78] S. Oono, I. Hitoshi, K. Ryoko, K. Daisuke, S. Masaki, and E. and Maeda, "Interaction strategy for robotic apology based on human orientation toward service," *Advanced Robotics,* vol. 38, no. 4, pp. 226-245, 2024.

[79] R. Casidy, and H. Shin, "The effects of harm directions and service recovery strategies on customer forgiveness and negative word-of-mouth intentions," *Journal of Retailing and Consumer Services,* vol. 27, pp. 103-112, 2015.

[80] B. F. Malle, and D. Ullman, "A multi-dimensional conception and measure of human-robot trust," *Trust in Human-Robot Interaction*, pp. 3-25, 2021.

[81] M. Heerink, K. Ben, V. Evers, and B. Wielinga, "The influence of social presence on acceptance of a companion robot by older people," *Journal of Physical Agents,* vol. 2, no. 2, pp. 33-40, 2008.

[82] J. S. Downs, M. B. Holbrook, S. Sheng, and L. F. Cranor, "Are your participants gaming the system? screening mechanical turk workers," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, Georgia, USA, pp. 2399–2402, 2010.

[83] D. M. Oppenheimer, T. Meyvis, and N. Davidenko, “Instructional manipulation checks: Detecting satisficing to increase statistical power,” *Journal of experimental social psychology,* vol. 45, no. 4, pp. 867-872, 2009.

[84] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, “G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences,” *Behavior research methods,* vol. 39, no. 2, pp. 175-191, 2007.

[85] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, “The aligned rank transform for nonparametric factorial analyses using only anova procedures,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, pp. 143–146, 2011.

[86] Y. Takano, and S. Sogon, “Are Japanese more collectivistic than Americans? Examining conformity in in-groups and the reference-group effect,” *Journal of Cross-Cultural Psychology,* vol. 39, no. 3, pp. 237-250, 2008.

[87] Y. Ohtsubo, E. Watanabe, J. Kim, J. T. Kulas, H. Muluk, G. Nazar, F. Wang, and J. Zhang, “Are costly apologies universally perceived as being sincere?: A test of the costly apology-perceived sincerity relationship in seven countries,” *Journal of Evolutionary Psychology JEP,* vol. 10, no. 4, pp. 187-204, 2012.

[88] D. Sakamoto, T. Kanda, T. Ono, H. Ishiguro, and N. Hagita, “Android as a telecommunication medium with a human-like presence,” in 2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 193-200, 2007.

[89] D. F. Glas, T. Minato, C. T. Ishi, T. Kawahara, and H. Ishiguro, “Erica: The erato intelligent conversational android,” in Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on, New York, NY, United States, pp. 22-29, 2016.

[90] M. Shiomi, H. Sumioka, K. Sakai, T. Funayama, and T. Minato, “SŌTO: An Android Platform with a Masculine Appearance for Social Touch Interaction,” in Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, Cambridge, United Kingdom, pp. 447–449, 2020.

[91] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu, “Development and evaluation of an interactive humanoid robot "Robovie",” in Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292), Washington, DC, United States, pp. 1848-1855 vol.2, 2002.

[92] S. Shamsuddin, L. I. Ismail, H. Yussof, N. I. Zahari, S. Bahari, H. Hashim, and A. Jaffar, “Humanoid robot NAO: Review of control and motion exploration,” in 2011 IEEE International Conference on Control System, Computing and Engineering, pp. 511-516, 2011.

[93] J. Bohren, R. B. Rusu, E. G. Jones, E. Marder-Eppstein, C. Pantofaru, M. Wise, L. Mösenlechner, W. Meeussen, and S. Holzer, “Towards autonomous robotic butlers: Lessons learned with the PR2,” in 2011 IEEE International Conference on Robotics and Automation, pp. 5568-5575, 2011.