This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

Designing Standing Position and Voice Cues for a Robot Riding in a Social Elevator

Masahiro Shiomi[a]*, Masayuki Kakio[b], and Takahiro Miyashita[a]

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

[b]*Mitsubishi Electric Corporation, Tokyo, 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.

Masayuki Kakio received M. Eng. and Ph.D. degrees in engineering from Osaka University in 2006 and 2011, respectively. He was an intern researcher at the Advanced Telecommunications Research Institute International (ATR). He joined Mitsubishi Electric Corporation in 2009. He was also a visiting scholar at University of Toronto (UofT) from 2013 to 2014. He is currently a senior manager in the Reliable Robotics section at the Robotics department. He is interested in the development of intelligent systems.

Takahiro Miyashita received his Ph.D. in engineering for computer-controlled machinery from Osaka University, Japan, in 2002. Currently, he serves as the Director of Interaction Science Laboratories (ISL) and Interaction Technology Bank (ITB) at Advanced Telecommunications Research Institute International (ATR). He is also one of group leaders of Moonshot R&D Program Goal 1 Avatar-Symbiotic Society Project under the supervision of Prof. Hiroshi Ishiguro at Osaka University. His research focuses on Cybernetic Avatars (CA), their communication infrastructure (CA-PF), human-robot interaction, and cloud-networked robotics. Beyond research, he is committed to fostering entrepreneurship and promoting industry-academia collaboration for the societal implementation of advanced technologies. In 2015, he established the i-RooBO Network Forum, a consortium of robotics and IoT companies. In 2021, he also established the Corporate Consortium for Avatar-Symbiotic Society (C-CAS2), a consortium specifically dedicated to companies involved in cybernetic avatar technologies.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

Designing Standing Position and Voice Cues for a Robot Riding in a Social Elevator

Recent technological advances have enabled robots to move across multiple floors by elevators. While various technologies allow robots to operate in such settings, most related studies have focused primarily on the robots themselves as they enter and ride elevators, e.g., examining the social behaviors of the robots around them. In contrast, limited attention has addressed the social behaviors from the elevator perspective, i.e., developing social elevators and their coordination with robots have received much less focus. To solve these problems, we first conducted a data collection to analyze passengers' preferred standing positions in an elevator to design standing positions for robots in social elevators. Building on the knowledge gleaned from the data collection, we decided on a mobile robot's behaviors in an elevator and investigated the effects of voice cues for both an elevator and a robot. In our experiment, either the elevator, the robot, or both provided voice cues to passengers when the robot rode the elevator and we evaluated the participant impressions of each device. The results showed that the robot's use of voice cues produced positive effects for the elevator and vice versa, although if both alternately deployed voice cues, no additional benefits were identified.

Keywords: human-elevator interaction; social elevator; robot-robot interaction; preferred standing position; voice cue

Introduction

Mobile robots have increasingly gained the ability to navigate buildings using elevators [1-4]. Several studies have enabled them to manipulate car operating panels by robotic arms [5-7], and others have employed vision-based methods for car operating panel recognition [8-10]. An alternative strategy is to upgrade the elevators themselves, allowing direct communication between elevators and mobile robots [11-13]. Through these approaches, researchers have demonstrated how robots can satisfy everyday mobility requirements in multi-floor environments.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

In addition to these technical contributions, some researchers have examined robots' social behaviors around elevators. For instance, studies have explored how robots might use visual and auditory cues (e.g., LED indicators or sound signals) when entering or riding an elevator with humans [14, 15]. Others investigated passengers' preferences regarding where a mobile robot or a person should stand inside the elevator [16], as well as whether robots should wait in a machine-like or a human-like manner [17]. Such findings indicate that social acceptability is critical when mobile robots and humans share elevator space.

Unfortunately, previous work has often focused on the robot's actions alone, such as how a robot itself signals intentions or communicates with passengers. In contrast, scant attention has addressed the elevator's perspective, particularly the potential for a social elevator that can convey voice cues or facilitate collaboration with mobile robots. We focused on this issue because some future low-cost robots may lack speech capabilities, requiring them to rely on the elevator's voice outputs to enhance social interactions. The elevator itself must directly interact with passengers. Additionally, if improved coordination between elevators and mobile robots increases the perceived appropriateness of voice cues, then both elevator technology and mobile robot design will benefit from new insights into collaborative behaviors.

Based on these considerations, we examined how the voice cue designs provided by an elevator and a mobile robot influence social acceptance when the latter is taking an elevator (Fig. 1). We investigated these questions by preparing a teleoperated mobile robot and a social elevator, each capable of audibly providing information to human passengers.

We additionally investigated the passengers' standing positions inside the elevator in actual settings. Although past studies reported the preferred standing positions of

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

passengers inside an elevator, one study was conducted on a web survey with an elevator of a specific size (in a different country from our experiments) [16] or using an elevator whose car operating panel was located only in the center of one side [18]. Preferred standing positions in elevators would be influenced by factors such as elevator design variations across different cultures and whether a country follows left- or right-hand traffic conventions, but these past studies have not fully addressed how these factors impact positioning in real-world elevator scenarios. Therefore, we reproduced a similar data collection but in actual elevator use cases to design more appropriate standing positions for mobile robots in Japan.

Following these considerations, we answer two research questions: **RQ1:** Where are the preferred standing positions for passengers in an elevator? **RQ2:** How should the voice cues be designed in combination with a social elevator and a mobile robot? Although this paper is an extended version of previous work by Shiomi et al. [19], it contains more related works, additional experiments and analyses that investigated standing positions in elevator use, and more detailed discussions.

Fig. 1 Mobile robot takes a social elevator

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

Related Work

Human-elevator interaction

A number of studies have examined how people use elevators, focusing on such aspects as in-elevator behavior, standing positions, and the role of waiting times. For instance, a past study employed a smart elevator that was equipped with multiple sensor devices to analyze standing preferences and movement patterns [18]. Another study investigated through a web-based survey where mobile robots and humans should stand in an elevator to ensure socially acceptable robot behaviors [16]. Another past study explored both waiting and transit times and perceived stress through online surveys, and their result emphasized the importance of psychological factors in elevator use [20]. From another perspective, several researchers achieved voice-based interaction capabilities for elevators [21, 22].

Collectively, these investigations unveiled essential elements of human-elevator interaction, including standing-position choices and the influence of waiting times. However, these standing-position studies were carried out in right-hand traffic countries, whereas our research is situated in Japan (a left-hand traffic context), where cultural or spatial preferences may differ. Understanding passengers' preferred standing locations is vital, particularly when determining where a robot should position itself. Therefore, we conducted our own data collection to capture local standing-position tendencies to provide insights into human-elevator interaction within our setting.

Robot-robot (elevator) interaction

In this study, we conceptualized the elevator itself as a robot with speaking capabilities (i.e., a social elevator) and investigated the effects of its “speaker role” alongside a mobile robot’s voice interaction. Prior work on multi-robot interactions informs our approach: various studies have examined how multiple robots perform such

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

tasks as information delivery [23, 24], educational and motivational support [25-27], persuasion [28-30], apologies [31, 32], recommendations [33-35], entertainment (e.g., comedy) [36, 37], and the expression of cute behaviors [38, 39]. These projects generally underscore the utility of deploying multiple robots and often categorize various styles of multi-robot communication. Moreover, in the context of robot-elevator interaction, robotics researchers have developed several functions, including vision-based detection of car operating panels [5-7], button-pressing capabilities by robotic arms [8-10], communication protocols between mobile robots and elevators [11-13], as well as identifying social behaviors when the robots shared an elevator with human passengers [14-17].

However, most such studies focus on conversational scenarios, overlooking situations where a mobile robot must take an elevator and how passengers perceive any collaboration between the robot and the elevator. By extending multi-robot voice interaction concepts to elevator usage, our research explores how passengers respond to a coordinated use of vocal cues by both a mobile robot and the social elevator itself, offering new insights into collaborative social behaviors in the context of elevator use.

Data collection

Overview

Identifying and comprehending the preferred standing position of passengers in elevators is critical for deciding where a mobile robot should stand when using elevators. If only a robot is taking the elevator, it should avoid places preferred by human passengers in consideration of the possibility that someone else will board it later. If others are already in the elevator, it should stand away from them.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

However, past studies reported an opposite phenomenon about preferred standing positions in elevators. One study concluded that passengers' preferred standing positions are near the car operating panel [18], although another reported that the preferred standing positions are in the back corners--away from the car operating panel [16]. Moreover, these studies were conducted in right-hand traffic countries, not left-hand traffic countries where our study is conducted. Another difference between these studies and ours is the number of car operating panels in the elevator; past studies' elevators have only one car operating panel; our elevator has two car operating panels (Fig. 1). Therefore, directly applying past knowledge about preferred standing positions in elevators is difficult for our experiment settings.

Based on these considerations, we collected data to investigate the preferred standing positions of passengers in elevators from two different situations: 1) only one passenger is using the elevator, and 2) two passengers are using it (i.e., one passenger already in the elevator is joined by another).

Participants

Thirty individuals (15 females and 15 males), all native Japanese speakers, took part in this study. Their average age was 34.2 years (SD = 9.00). They were recruited through a local commercial agency in Japan and were compensated (about $15) for their participation. Each wore a jacket that prominently displayed an identification number to facilitate data collection.

Environment

Figure 2 illustrates the experiment's setting. The study took place in a laboratory elevator equipped with car operating panels on both sides of the door. The building has four floors, including one basement level and three above-ground levels. The elevator

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at:
https://doi.org/10.1080/01691864.2025.2508782

interior was approximately 2.3 m high, 1.35 m wide, and 1.6 m deep. We installed video cameras inside it and near the ceiling of the adjacent hallway to record the proceedings.

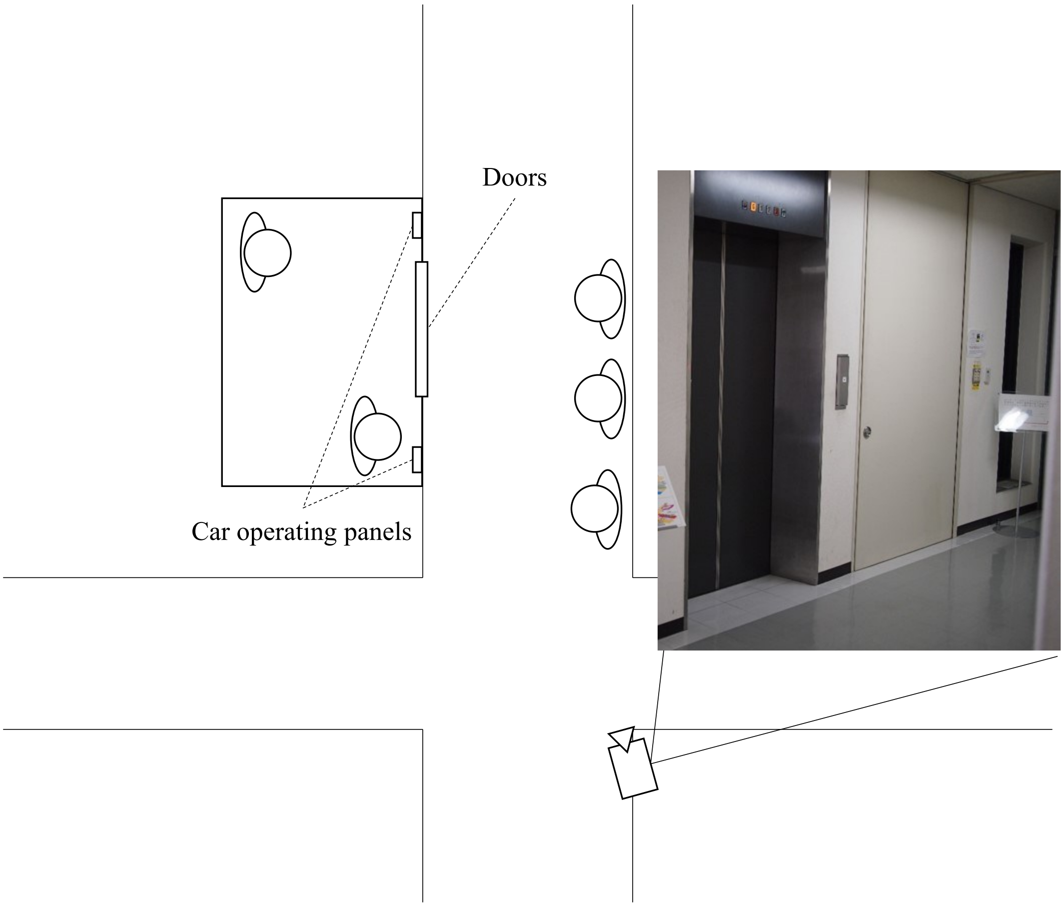


Fig. 2 Data collection environment

Measurement

We measured the standing positions of the passengers using video images. We divided the standing positions in the elevator into six sections (Fig. 3). The position where participants remained the longest during their elevator ride was identified as their standing position.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

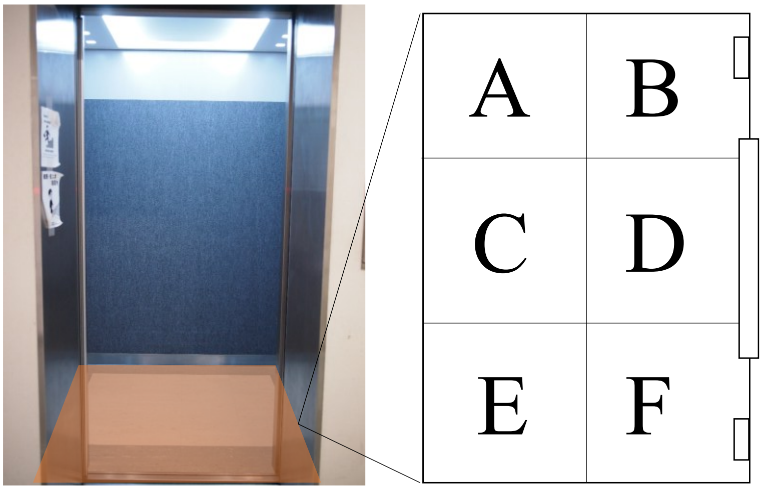


Fig. 3 Standing-position categories

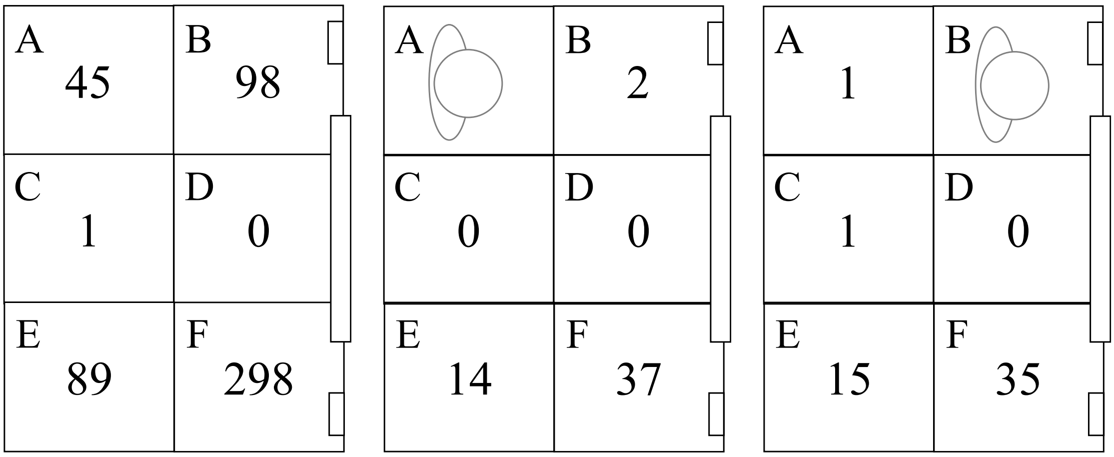
Procedure

All the procedures were approved by the Advanced Telecommunications Research Institute International Review Board Ethics Committee (525H). To replicate a range of real-world elevator conditions, multiple participants took part concurrently. However, to prevent undue congestion inside the elevator at any one time, the group size was limited to five participants. Each initially boarded the elevator from the third floor and followed a pre-assigned sequence: stopping at a designated floor and getting off, waiting for the next elevator, and traveling to another floor. Each participant rode the elevator 48 times.

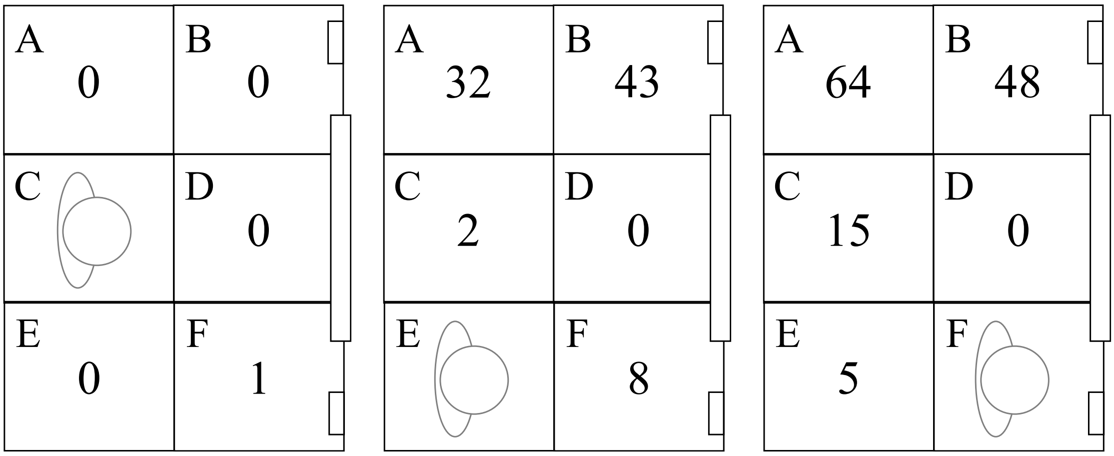
Data collection results

Based on the video images installed in the elevator, our analysis results identified 1152 valid data: 531 data when only one passenger used the elevator and 323 data when a passenger was already in it. These two data types accounted for about 70% of the total. In this paper, we only analyzed these data to investigate the preferred standing positions in the elevator (Fig. 4).

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782



(a) only one passenger (b) a passenger in A (c) a passenger in B



(d) a passenger in C (e) a passenger in E (f) a passenger in F

Fig. 4 Preferred standing positions in elevator

Only one passenger in the elevator

Figure 4-a shows the numbers and ratios of the standing positions of the participants in the elevator when only one participant was riding it. Many of the participants preferred the front of the left-side, car operating panel (position F). The left-back corner (position E) and the front of the right-side, car operating panel (position B) showed similar numbers. Participants generally avoided being directly in front of the doors (positions C/D). No participant chose D, which is directly in the front of the door,

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

and only one participant chose position C, which is also in the front of the door, although slightly toward the back of the car.

Two passengers in the elevator

Figures 4-b to f show the numbers and ratios of the standing positions of the participants in the elevator when one is already in it and another enters. These results showed that participants preferred an opposite side position from the participants who were already in the elevator. Only when participants were in position B did the majority stand near the opposite car operating panel rather than in the diagonal position.

Implication and standing-position design for mobile robots

Our analysis showed that the participants in our experiments preferred to stand on the left-side, car operating panels when the elevator was empty. This phenomenon contradicts a past study, which reported that passengers preferred to take standing positions in the right-rear of the corners [16]; our study resembles another past study [18], which also reported that passengers preferred standing positions near the car operating panel. Another interesting point is that our participants preferred the left side of the elevator, even though it has car operating panels on both sides. Since this past study was conducted in right-hand traffic countries [16], where people generally stand on the right side, perhaps people in left-hand traffic countries prefer to stand on the left-hand side.

Moreover, the participants preferred to stand diagonally opposite the previous person, in other words, maintaining as much distance as possible. This phenomenon is also similar to past studies [16, 18]. Following these results, the appropriate standing position for a mobile robot in the elevator in our study is the right-rear position. Even if

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

a passenger gets on an elevator that already has a robot, the passenger can take the left-front position where most people choose to stand.

Experiment

Our experiment investigated the effects of voice cues for both a mobile robot and a social elevator. We employed the above data collection knowledge to decide the standing positions of the participants and the robot to control the experiment settings. Detailed information is given below.

Hypothesis and prediction

Prior research suggests that mobile robots' social behaviors, such as voice cues, can facilitate acceptable interactions with passengers when traveling by elevator [14, 15]. Recent smart elevators also offer passenger-facing interaction capabilities [40, 41], although earlier investigations did not adequately address such social behaviors as elevator voice cues or collaborative interactions between elevators and robots. Other work has highlighted the effectiveness of employing multiple robots in conversational contexts [25-27]. Given the limited exploration of how elevators and robots might cooperate through voice-based interactions, we hypothesize that such collaborative behaviors will positively impact passenger impressions. Based on this perspective, we proposed the following predictions:

Prediction 1: Participants will view the robot more favorably when it provides voice cues than when it remains silent.

Prediction 2: Participants will view the elevator more favorably when it provides voice cues than when it remains silent.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

- **Prediction 3:** Participants will view both the elevator and the robot more favorably when both speak, compared to situations where only one or neither uses voice cues.

Participants

Thirty individuals (15 women and 15 men) participated in this study. They did not participate in the data collection. Their ages ranged from 20s to 50s, with an average of 39.4 years (SD = 11.6). They were recruited through the same temporary employment agency of the data collection and were compensated (about $15) for their participation.

Environment

Figure 5 provides an overview of the experimental setting. The study took place in our laboratory using a mobile robot (Fig. 6, left) and an elevator (Fig. 6, right). The next subsection contains additional details on the former's specifications. The elevator (2.3 m high, 1.35 m wide, and 1.6 m deep) featured car operating panels on both sides of its doorway (Fig. 1). For safety, the robot was controlled from an operator area enclosed by partitions. A network camera was installed inside the elevator for teleoperation.

This is an Accepted Manuscript of an article published by Taylor & Francis in
Advanced Robotics on 09 Mar 2025, available at:
https://doi.org/10.1080/01691864.2025.2508782

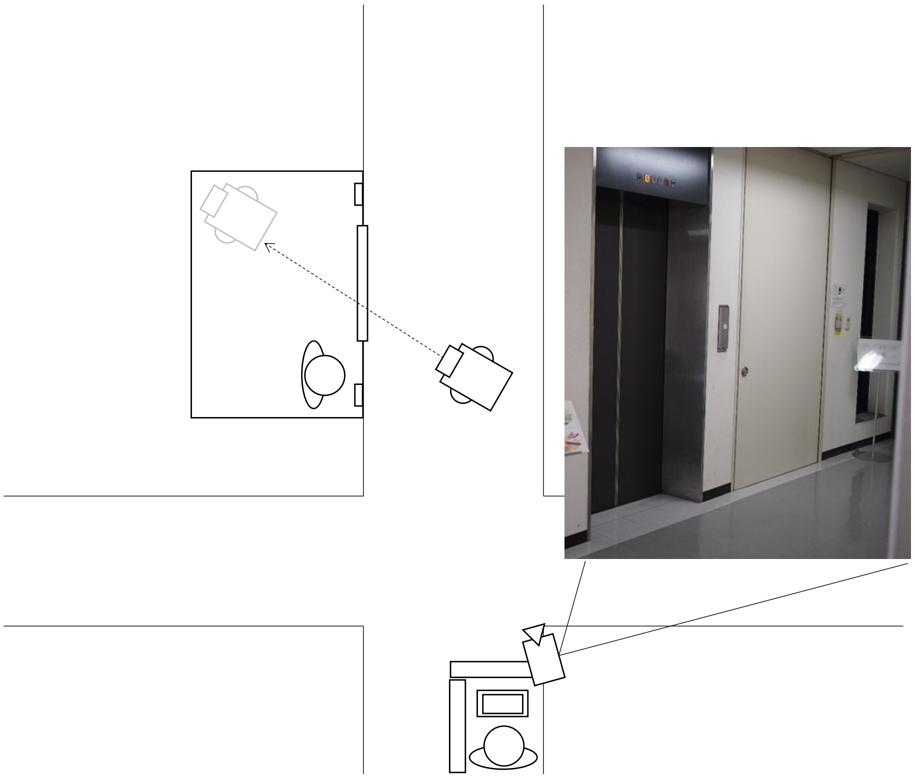


Fig. 5 Experimental environment

Fig. 6 Robot (left) and interior of social elevator (right)

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

Robots and teleoperation system

We employed a mobile robot (1,420 mm tall, 500 mm wide, and 630 mm deep) that also included a shelf on a mobile base. Its maximum moving speed in the elevator was 160 mm/s. Its display presented two eye-like features resembling a face, and a speaker was mounted on the top of the shelf.

We used a wizard-of-oz (WoZ) teleoperation system [42] for control. An operator managed the robot's locomotion speed, its speech function (VOICEVOX, KotoyomiNia), and the elevator's speech output (VOICEVOX, NekotukaR).

Conditions

We established four experimental conditions (Fig. 7). We varied two factors, i.e., the *elevator-speech* factor (none or present) and the *robot-speech* factor (none or present). When either the elevator or the robot spoke, it delivered two sentences: "The robot will now take the elevator" and "Please wait a moment." These phrases were deliberately chosen to offer simple and clear explanations of the robot's actions through discussions among the authors based on existing announcements used in commercial elevators and mobile robot systems in order to communicate the systems' intentions to nearby users in a manner that is easily understandable. In the condition when both the elevator and the robot spoke, each provided one sentence, maintaining an equivalent total amount of information. When both the elevator and the robot spoke (Fig. 7-D), due to the prepared speech contents, the elevator spoke first, followed by the robot.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

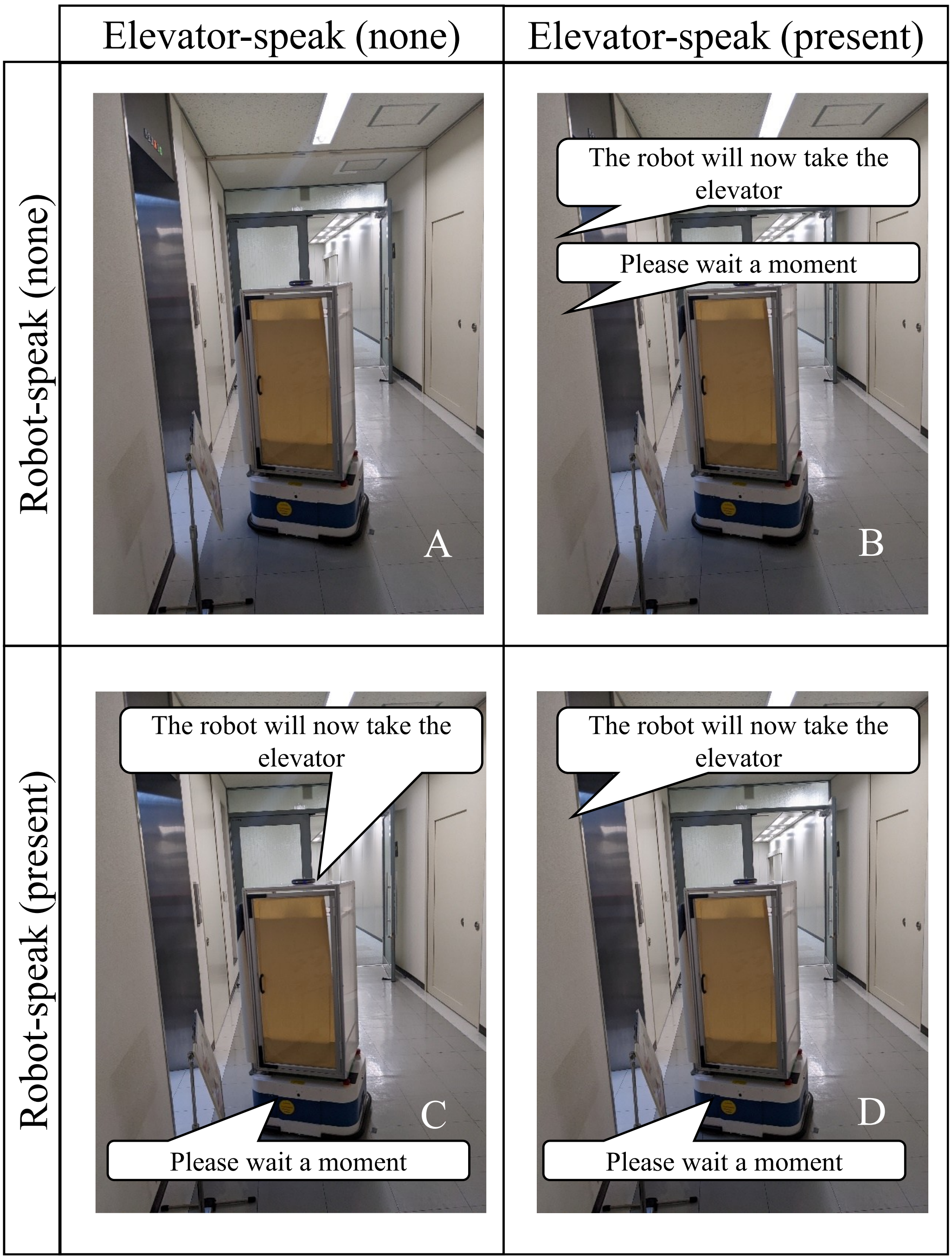


Fig. 7 Illustration of experiment conditions

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

Fig. 8 Robot takes elevator while participant waits inside

Measurements

We used three established questionnaire scales to evaluate the participants' impressions of both the elevator and the robot: likeability, perceived intelligence, and safety [43]. Each item was rated on a 7-point scale, with 1 indicating the least favorable response and 7 the most favorable.

Procedures

All the procedures were approved by the Advanced Telecommunications Research Institute International Review Board Ethics Committee (525H). Participants first read written instructions explaining the experiment and how to assess the elevator and robot in each condition. A within-participant design was employed, allowing each

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

participant to experience all four conditions in a counterbalanced order. Participants rode the elevator first, followed by the robot (Fig. 8). Based on the data collection results and maintaining consistency among the participants, they stood in front of the left car operating panel in every condition. Depending on the condition, just one or both the elevator and the robot spoke as the robot approached to enter the elevator. Inside the elevator, the robot was positioned in the right-rear corner, on the side opposite of the participant. Once the robot completed its ride, participants answered questionnaires.

Experiment results

Questionnaire results

We performed a two-factor ANOVA on each questionnaire scale (likeability, intelligence, safety) for both the elevator and the robot. The following are the two between-participant factors: (1) *elevator-speech* (none or present) and (2) *robot-speech* (none or present). Below, we report the F-tests, the *p*-values, and the partial eta-squared (η^2) effect sizes, as well as relevant simple main effects.

Regarding the elevator's likeability (Fig. 9-a), we found a significant effect in the *elevator-speech* factor ($F(1, 29) = 10.768$, $p = 0.003$, *partial* $\eta^2 = 0.271$), in the *robot-speech* factor ($F(1, 29) = 5.751$, $p = 0.023$, *partial* $\eta^2 = 0.165$), and in the interaction effects ($F(1, 29) = 19.845$, $p < 0.001$, *partial* $\eta^2 = 0.406$). For the simple main effects, when the elevator did not speak, participants rated the elevator significantly more positively if the robot spoke than if it did not ($p < 0.001$). When the robot did not speak, participants rated the elevator significantly more positively if the elevator spoke than if it did not ($p < 0.001$).

Regarding the elevator's intelligence (Fig. 9-b), we found a significant effect in the *elevator-speech* factor ($F(1, 29) = 32.235$, $p < 0.001$, *partial* $\eta^2 = 0.526$) and in the

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

interaction effects ($F(1, 29) = 16.230$, $p < 0.001$, *partial* $\eta^2 = 0.359$). We did not find a significant effect in the *robot-speech* factor ($F(1, 29) = 1.670$, $p = 0.206$, *partial* $\eta^2 = 0.054$). Concerning the simple main effects, when the elevator did not speak, participants rated the elevator significantly more positively if the robot spoke than if it did not ($p = 0.004$). When the robot did not speak, participants rated the elevator significantly more positively if the elevator spoke than if it did not ($p < 0.001$).

Regarding the elevator's safety (Fig. 9-c), we found a significant effect in the interaction effects ($F(1, 29) = 5.351$, $p = 0.028$, *partial* $\eta^2 = 0.156$). We did not find a significant effect in the *elevator-speech* factor ($F(1, 29) = 3.061$, $p = 0.091$, *partial* $\eta^2 = 0.095$) or in the *robot-speech* factor ($F(1, 29) = 0.364$, $p = 0.551$, *partial* $\eta^2 = 0.012$). Concerning the simple main effects, when the robot did not speak, the participants rated the elevator significantly more positively if the elevator spoke than if it did not ($p = 0.007$).

Regarding the robot's likeability (Fig.9-d), we found a significant effect in the *robot-speech* factor ($F(1, 29) = 46.991$, $p < 0.001$, *partial* $\eta^2 = 0.618$) and in the interaction effects ($F(1, 29) = 32.977$, $p < 0.001$, *partial* $\eta^2 = 0.532$). We did not find a significant effect in the *elevator-speech* factor ($F(1, 29) = 2.739$, $p = 0.109$, *partial* $\eta^2 = 0.086$). For the simple main effects, when the elevator did not speak, participants rated the robot significantly more positively if the robot spoke than if it did not ($p < 0.001$). When the elevator spoke, participants rated the robot significantly more positively if the robot spoke than if it did not ($p = 0.040$). When the robot did not speak, participants rated the robot significantly more positively if the elevator spoke than if it did not ($p < 0.001$).

Regarding the robot's intelligence (Fig. 9-e), we found a significant effect in the *elevator-speech* factor ($F(1, 29) = 4.838$, $p = 0.036$, *partial* $\eta^2 = 0.143$), in the *robot-*

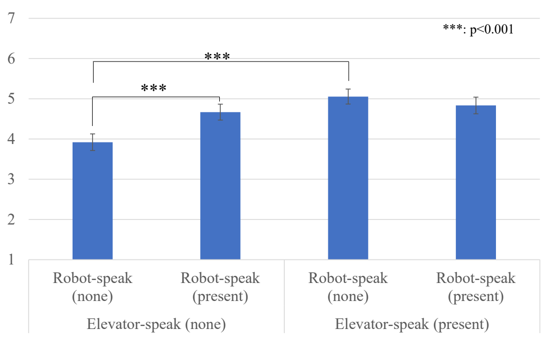
This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

speech factor ($F(1, 29) = 18.512$, $p < 0.001$, *partial* $\eta^2 = 0.390$), and in the interaction effects ($F(1, 29) = 18.560$, $p < 0.001$, *partial* $\eta^2 = 0.390$). For the simple main effects, when the elevator did not speak, participants rated the robot significantly more positively if the robot spoke than if it did not ($p < 0.001$). When the robot did not speak, participants rated the robot significantly more positively if the elevator spoke than if it did not ($p < 0.001$).

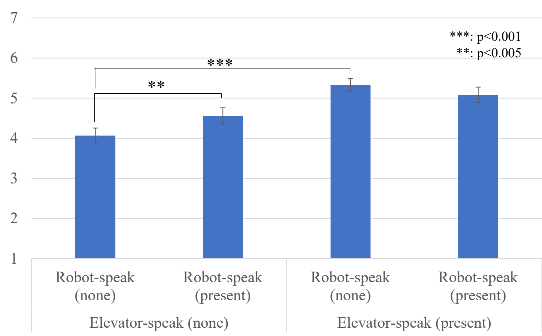
Regarding the robot's safety (Fig. 9-f), we did not found a significant effect in the *elevator-speech* factor ($F(1, 29) = 0.972$, $p = 0.332$, *partial* $\eta^2 = 0.032$), in the *robot-speech* factor ($F(1, 29) = 0.833$, $p = 0.369$, *partial* $\eta^2 = 0.028$), or in the interaction effects ($F(1, 29) = 0.207$, $p = 0.652$, *partial* $\eta^2 = 0.007$).

Summary of analysis results

Overall, Predictions 1 and 2 are partially supported, since the participants rated the elevator and the robot more positively (on likeability and intelligence measures) when they employed voice cues, although this effect did not extend consistently to safety. Prediction 3 was not supported: the participants did not exhibit significantly more positive impressions when both the elevator and robot spoke relative to scenarios in which only one or the other used voice cues.

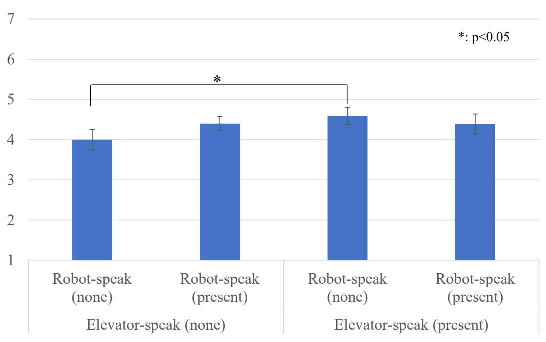


(a) elevator's likeability

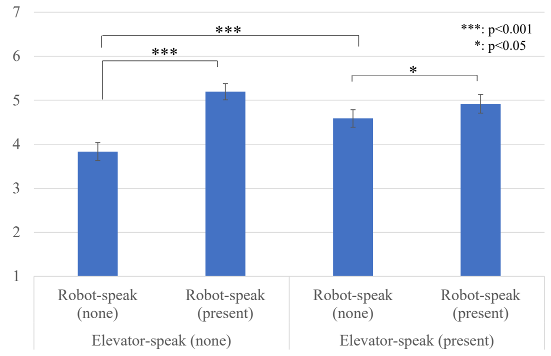


(b) elevator's intelligence

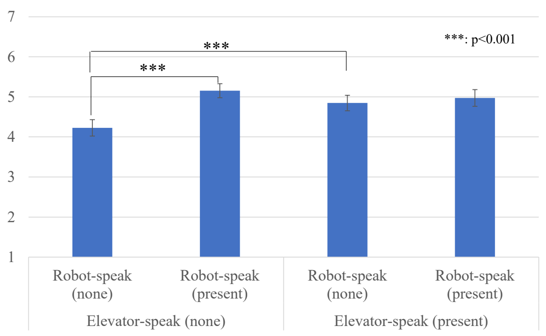
This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782



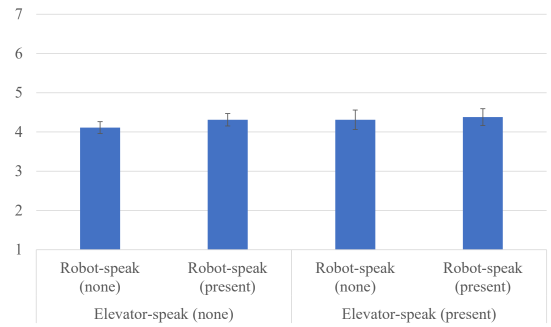
(c) elevator's safety



(d) robot's likeability



(e) robot's intelligence



(f) robot's safety

Fig. 9 Questionnaire results (displayed simple main effects only)

Implications from experiment results

Our findings indicate that voice cues enhanced the participants' impressions of both the social elevator and the mobile robot, except for the latter's perceived safety, even when only one of them produced voice output. This phenomenon is somewhat different from earlier work on interactions among multiple robots [44], possibly due to the distinctive context of elevator use rather than a straightforward conversational setting. Whereas prior studies emphasized dialog-based interactions between robots and participants, our research concentrated on a realistic scenario where a person and a mobile robot share an elevator space.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

One interesting observation is the complementary influence of voice cues on perceived likeability and intelligence. When the elevator provided the voice cues, participants rated the elevator itself and the robot more positively in terms of likeability and intelligence. Conversely, when the robot provided the voice cues, participants perceived improvements in the likeability of both entities, yet the elevator's perceived intelligence remained unaffected.

A possible interpretation for this asymmetry is related to how participants socially construed the interaction scenario. Because robots are generally perceived as social actors capable of intentional communication, the elevator’s voice cues may have suggested greater sophistication or capability in the robot, as if the robot were interacting with an advanced, communicative environment. On the other hand, when the robot itself provided voice cues, participants might have interpreted the elevator merely as a passive tool being operated or controlled by the robot rather than as a social elevator, i.e., a sophisticated or intelligent entity in its own right.

This interpretation implies that the social role participants attributed to each entity influenced their perceptions. Elevators, typically viewed as tools, become perceived as socially interactive entities primarily when they actively initiate interaction. Conversely, robots, already perceived as socially interactive by default, may reinforce their social agency through verbal interactions initiated by other entities in the environment.

Such asymmetry implies that merely having two agents present, where one of them speaks, may be sufficient to change the perceived attributes of the other. Although previous multi-robot research generally involved two robots that both contributed comparable amounts of verbal content [44], a recent study reported that information-providing with imbalanced verbal content between two robots remains more effective

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

than information-providing by a single robot [39]. Our study might suggest how imbalanced interactions in robot-robot interaction changed people's impressions.

Generalizability and future directions regarding voice cue effects

In this study, we used specific verbal phrases as cues, intentionally designed to represent typical announcements used in real-world robotic and elevator systems. The primary function of these verbal cues is to communicate the system's current status and upcoming actions clearly to users. Therefore, minor variations in phrasing, as long as they fulfill this communicative role, are unlikely to alter user impressions significantly.

Nevertheless, future research could explore more elaborate speech strategies, such as providing additional informative content to alleviate user stress during waiting periods. Another promising direction is investigating how variations in voice characteristics (e.g., gender, tone, speaking style) affect user perceptions and experiences. Although these questions extend beyond the scope of the present study, addressing them would offer valuable insights for the practical deployment of social elevators and mobile robots in real-world environments.

General Discussion

Implication

Our data collection results provide an intriguing implication about preferred standing positions in elevators because our study reported such positions surveyed in a left-traffic country, which is different from past studies that investigated preferred standing positions in right-traffic countries. Although such elevator characteristics as the size, shape, and number of car operating panels differ among these studies, our study provides useful knowledge in the context of not only human-elevator interaction but also human-robot and robot-elevator interaction contexts.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

Moreover, our experiment's insights could benefit researchers and developers of next-generation elevator systems. Earlier research on robot-elevator interaction has predominantly focused on the robot’s social behavior when sharing an elevator with human passengers [16, 17], leaving the elevator itself relatively unexamined. Yet in everyday life, many individuals regularly ride elevators, positioning them as potential “social robots” whose behaviors should be designed for acceptability. In practice, a social elevator might coordinate with a robot in advance, conveying information to passengers during their ride. However, simpler commercial robots often lack speech capabilities, suggesting that an elevator’s capacity to speak could meaningfully improve user acceptance of co-riding robots.

In addition, note that voice cues did not significantly affect the perception of safety, perhaps because the mobile robot's standing positions were based on our data collection. The participants might feel safe on the mobile robot because they stand the diagonal on opposite sides, as humans do. Additional social behaviors, such as the robot’s movement path, could be investigated to address safety-related perceptions.

Limitation

Several limitations exist in our study. First, in the data collection, we employed only a specific elevator in a specific country, i.e., Japan. Therefore, our data collection setting cannot cover situations where passengers use different types of elevators in various countries. Next, in the experiment, we employed only a single mobile robot and only one specific elevator in a particular country, too. Additionally, just one passenger in the experiment rode the elevator at a time. Future studies should examine different robot types (e.g., multi-legged or humanoid forms [45, 46]) and various conditions, such as the number of passengers. Another limitation is that the elevator always spoke first in

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

our design, potentially influencing how participants perceived subsequent speech. Despite these constraints, our work offers meaningful guidance to robotics researchers seeking design social interactions for robots and elevators.

Conclusion

In this study, we first investigated preferred standing positions in an elevator involving multiple participants. Our data collection results showed that the participants preferred to stand near the left-side, car operating panels when the elevator was empty, and they preferred to stand diagonally opposite the person with whom the elevator was being shared. We next investigated with a WoZ approach how voice cues from a mobile robot and a social elevator affected passenger impressions. We found that each device's voice cue enhanced positive impressions of the other, although no combined effect emerged when both spoke in turns. The elevator's voice cues bolstered the robot's perceived intelligence, and the robot's voice cues improved the elevator's perceived likeability. These asymmetric interaction effects offer important considerations for developers of both elevator and robot systems.

Acknowledgement

This work was partially supported by JST CREST (Core Research for Evolutional Science and Technology) Grant, Number JPMJCR18A1, Japan, JSPS KAKENHI Grant Numbers JP24K21327, and Mitsubishi Electric Corporation. We thank Kazutoshi Akazawa, Haruka Yamazaki, Takuya Murakami, Yuji Yamauchi, and Sayuri Yamauchi for their help during the execution of our experiments. We used ChatGPT (GPT-4o) to correct and phrase the English texts.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

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.

References

[1] J. Palacín, R. Bitriá, E. Rubies, and E. Clotet, "A Procedure for Taking a Remotely Controlled Elevator with an Autonomous Mobile Robot Based on 2D LIDAR," *Sensors,* vol. 23, no. 13, pp. 6089, 2023.

[2] J. Liebner, A. Scheidig, and H.-M. Gross, "Now i need help! passing doors and using elevators as an assistance requiring robot," in International Conference on Social Robotics, pp. 527-537, 2019.

[3] J. Huang, T. Lau, and M. Cakmak, "Design and evaluation of a rapid programming system for service robots," in 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 295-302, 2016.

[4] J. Collin, J. Pellikka, and J. T. J. Penttinen, "Elevator Industry: Optimizing Logistics on Construction Sites with Smart Elevators," *5G Innovations for Industry Transformation: Data-Driven Use Cases*, J. Collin, J. Pellikka and J. T. J. Penttinen, eds., pp. 141-155: John Wiley & Sons, 2023.

[5] A. A. Ali, M. M. Ali, N. Stoll, and K. Thurow, "Integration of navigation, vision, and arm manipulation towards elevator operation for laboratory transportation system using mobile robots," *Journal of Automation, Mobile Robotics and Intelligent Systems,* vol. 11, no. 4, pp. 34-50, 2017.

[6] H. Yu, L. Li, J. Chen, Y. Wang, Y. Wu, M. Li, H. Li, Z. Jiang, X. Liu, and T. Arai, "Mobile robot capable of crossing floors for library management," in 2019 IEEE international conference on mechatronics and automation (ICMA), pp. 2540-2545, 2019.

[7] D. Zhu, Z. Min, T. Zhou, T. Li, and M. Q.-H. Meng, "An autonomous eye-in-hand robotic system for elevator button operation based on deep recognition network," *IEEE Transactions on Instrumentation and Measurement,* vol. 70, pp. 1-13, 2020.

[8] E. Klingbeil, B. Carpenter, O. Russakovsky, and A. Y. Ng, "Autonomous operation of novel elevators for robot navigation," in 2010 IEEE International Conference on Robotics and Automation, pp. 751-758, 2010.

[9] D. Zhu, T. Li, D. Ho, T. Zhou, and M. Q. Meng, "A novel OCR-RCNN for elevator button recognition," in 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3626-3631, 2018.

[10] D. Zhu, Y. Fang, Z. Min, D. Ho, and M. Q.-H. Meng, "Ocr-rcnn: An accurate and efficient framework for elevator button recognition," *IEEE Transactions on Industrial Electronics,* vol. 69, no. 1, pp. 582-591, 2021.

[11] J. López, D. Pérez, E. Zalama, and J. Gómez-García-Bermejo, "Bellbot-a hotel assistant system using mobile robots," *International Journal of Advanced Robotic Systems,* vol. 10, no. 1, pp. 40, 2013.

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

[12] Panasonic. "Panasonic Autonomous Delivery Robots -HOSPI- Aid Hospital Operations at Changi General Hospital.," 05 December 2023; https://news.panasonic.com/global/topics/2015/44009.html (
[13] A. A. Abdulla, H. Liu, N. Stoll, and K. Thurow, “A secure automated elevator management system and pressure sensor based floor estimation for indoor mobile robot transportation,” *Advances in Science, Technology and Engineering Systems Journal,* vol. 2, no. 3, pp. 1599-1608, 2017.
[14] W.-t. Law, K.-s. Li, K.-w. Fan, T. Mo, and C.-k. Poon, “Friendly Elevator Co-rider: An HRI Approach for Robot-Elevator Interaction,” in 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 865-869, 2022.
[15] F. Babel, P. Hock, J. Kraus, and M. Baumann, “Human-Robot Conflict Resolution at an Elevator-The Effect of Robot Type, Request Politeness and Modality,” in 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 693-697, 2022.
[16] D. Gallo, S. Gonzalez-Jimenez, M. A. Grasso, C. Boulard, and T. Colombino, “Exploring Machine-like Behaviors for Socially Acceptable Robot Navigation in Elevators,” in 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 130-138, 2022.
[17] D. Gallo, P. L. Bioche, J. K. Willamowski, T. Colombino, S. Gonzalez-Jimenez, H. Poirier, and C. Boulard, “Investigating the Integration of Human-Like and Machine-Like Robot Behaviors in a Shared Elevator Scenario,” in Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction, Stockholm, Sweden, pp. 192–201, 2023.
[18] T. Robal, K. Basov, U. Reinsalu, and M. Leier, “A Study into elevator passenger in-cabin behaviour on a smart-elevator platform,” *Baltic Journal of Modern Computing,* vol. 10, no. 4, pp. 665-688, 2022.
[19] 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.
[20] C. Bird, R. Peters, E. Evans, and S. Gerstenmeyer, “Your Lift Journey–How Long Will You Wait?,” in 6th Symposium on Lift and Escalator Technologies, pp. 53-64, 2016.
[21] D. Meenatchi, R. Aishwarya, and A. Shahina, “A Voice Recognizing Elevator System,” in Proceedings of the International Conference on Soft Computing Systems, New Delhi, pp. 179-187, 2016.
[22] A. González-Docasal, J. Alonso, J. Olaizola, M. Mendicute, M. P. Franco, A. d. Pozo, D. Aguinaga, A. Álvarez, and E. Lleida, “Design and Evaluation of a Voice-Controlled Elevator System to Improve the Safety and Accessibility,” *IEEE Open Journal of the Industrial Electronics Society,* vol. 5, pp. 1239-1250, 2024.
[23] A. Dahiya, A. M. Aroyo, K. Dautenhahn, and S. L. Smith, “A survey of multi-agent Human–Robot Interaction systems,” *Robotics and Autonomous Systems,* vol. 161, pp. 104335, 2023.
[24] S. Sebo, B. Stoll, B. Scassellati, and M. F. Jung, “Robots in Groups and Teams: A Literature Review,” *Proc. ACM Hum.-Comput. Interact.,* vol. 4, no. CSCW2, pp. Article 176, 2020.
[25] H. Itahara, M. Kimoto, T. Iio, K. Shimohara, and M. Shiomi, “How Does Exposure to Changing Opinions or Reaffirmation Opinions Influence the

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

Thoughts of Observers and Their Trust in Robot Discussions?," *Applied Sciences,* vol. 13, no. 1, pp. 585, 2023.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

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

[37] J. Swaminathan, S. Jujjuri, and H. Knight, "Using Street-Performance Style Robot Comedians To Attract Audiences For HRI Studies," *Workshop on Performing Art Robots & Technologies (PAR-T) at International Conference of Robots and Systems*, 2020.

[38] 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.

[39] 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,"

This is an Accepted Manuscript of an article published by Taylor & Francis in Advanced Robotics on 09 Mar 2025, available at: https://doi.org/10.1080/01691864.2025.2508782

in 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN), pp. 485-490, 2024.
[40] D. Meenatchi, R. Aishwarya, and A. Shahina, "A voice recognizing elevator system," in Proceedings of the International Conference on Soft Computing Systems: ICSCS 2015, Volume 1, pp. 179-187, 2016.
[41] S. Gupta, S. Tyagi, and K. Kishor, "Study and development of self sanitizing smart elevator," *Proceedings of Data Analytics and Management: ICDAM 2021, Volume 1*, pp. 165-179: Springer, 2022.
[42] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, "Wizard of Oz studies: why and how," in Proceedings of the 1st international conference on Intelligent user interfaces, Orlando, Florida, USA, pp. 193-200, 1993.
[43] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International Journal of Social Robotics,* vol. 1, no. 1, pp. 71-81, 2009.
[44] 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.
[45] 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.
[46] 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, pp. 447-449, 2020.