



PDF Download
3765766.3765777.pdf
08 January 2026
Total Citations: 1
Total Downloads: 47

 Latest updates: <https://dl.acm.org/doi/10.1145/3765766.3765777>

RESEARCH-ARTICLE

One Robot, Many Minds: Factors Shaping Visitors' Evaluation of an Autonomous Museum Robot Guide

Published: 10 November 2025

[Citation in BibTeX format](#)

HAI '25: International Conference on Human-Agent Interaction
November 10 - 13, 2025
Yokohama, Japan

One Robot, Many Minds: Factors Shaping Visitors' Evaluation of an Autonomous Museum Robot Guide

Luca Garello*
Contact
Italian Institute of Technology
Genoa, Italy
Softbots
Italian Institute of Technology
Genoa, Italy
luca.garello@iit.it

Francesca Cocchella*
Contact
Italian Institute of Technology
Genoa, Italy
francesca.cocchella@iit.it

Manuel G. Catalano
SoftBots
Italian Institute of Technology
Genoa, Italy
manuel.catalano@iit.it

Alessandra Sciutti
Contact
Italian Institute of Technology
Genoa, Italy
alessandra.sciutti@iit.it

Francesco Rea
Contact
Italian Institute of Technology
Genoa, Italy
francesco.rea@iit.it

Abstract

Robots are no longer just tools—they are becoming social agents that can shape how we engage with culture. This study examines what influences visitors' perceptions of an autonomous museum guide robot, focusing not only on technical capabilities but also on human-centered factors. In a maritime exhibition, 34 participants interacted with a fully autonomous, LLM-powered robot acting as a museum guide. Using self-report questionnaires, we explored how individual differences - age and prior experience with robots – interacted with experimental conditions to shape participants' impressions of the robot. Our findings suggest that these personal factors significantly affect how visitors evaluate the robot, suggesting that effective design must reflect the diversity of users' experiences and expectations. By acknowledging the complexity of human-robot interaction, we move closer to creating robotic guides that are not only functional but also socially attuned.

CCS Concepts

• Human-centered computing → Empirical studies in HCI; Empirical studies in interaction design.

Keywords

human-robot interaction, large language models, museum, evaluation

ACM Reference Format:

Luca Garello, Francesca Cocchella, Manuel G. Catalano, Alessandra Sciutti, and Francesco Rea. 2025. One Robot, Many Minds: Factors Shaping Visitors' Evaluation of an Autonomous Museum Robot Guide. In *13th International Conference on Human-Agent Interaction (HAI '25)*, November 10–13, 2025.

*Both authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution 4.0 International License. *HAI '25, Yokohama, Japan*

© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2178-6/25/11
<https://doi.org/10.1145/3765766.3765777>

Yokohama, Japan. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3765766.3765777>



Figure 1: Participant visiting the exhibition with the robot Alter-Ego.

1 Introduction

Robotic agents can be integrated into public and cultural spaces like museums, acting not just as technology but as social actors engaging visitors. Museum guide robots can enhance experiences by providing personalized information and encouraging interaction with exhibits. In Human-Robot Interaction (HRI), understanding perceptions during first encounters is key, requiring attention not only to robot capabilities but also to visitor diversity. Integrating verbal and spatial interaction—linking dialogue with navigation—is especially important, while demographic characteristics strongly shape attitudes toward autonomous systems in heterogeneous public contexts like museums.

Recent advancements in artificial intelligence, large language models (LLMs), and mobile robotics have enabled the development of autonomous systems capable of naturalistic dialogue and intelligent navigation in complex indoor spaces. Building on this progress, in our previous work [21] we presented a humanoid robot equipped

with autonomous navigation via SLAM, and an LLM-powered dialogue system. The robot was deployed in a maritime museum, where it guided visitors through a semi-structured exhibition tour (Figure 1). In that study, we evaluated how participants' perceptions of the robot changed from pre- to post-interaction and preliminarily explored differences between the two experimental conditions. In one, the interaction with the robot was announced in advance (Announced condition), in another, the interaction occurred unexpectedly (Surprised condition). While the initial analysis focused on overall changes in perception and the impact of context, it did not fully explore how individual factors, such as age and participants' prior experience with robots, might moderate these effects. Yet, previous works in HRI (e.g., [32]) suggest that user characteristics significantly shape robot evaluations, particularly in early encounters with novel systems. Importantly, these individual differences may interact with how the encounter is framed.

In this study, we focus on age and prior experience with robots as two key factors that may moderate how visitors evaluate an autonomous museum guide. This work goes beyond evaluating interactional outcomes to examine how people construct meaning around autonomous systems in real-world contexts. By focusing on first encounters in a cultural setting, it highlights both the dynamics of human-agent interaction and the processes through which users interpret the robot's role and presence.

We conducted our study **in the wild**, stressing the need for more HRI research in real-world environments such as workplaces, homes, and public spaces, and to expand the still limited understanding of how people engage with robots in complex social contexts [26][9]. The present paper extends the previous study by examining how stakeholders' individual characteristics, influence the effect of the experimental condition on their evaluations of the robot.

By explicitly examining the interplay between **framing conditions** and **individual differences**, our study contributes new insights into how **autonomous robots** are evaluated in public spaces, which can inform the design of robots that better adapt to diverse users in cultural venues. This contribution shares the procedure with our previous work [21], which we briefly summarize here for completeness. This paper provides entirely new statistical analyses, new research questions, and a broader interpretation.

2 Background

Autonomous robotic systems have progressively gained relevance in museum contexts, with early prototypes such as RoboX [36] and CiceRobot [11] demonstrating the potential of robots to support cultural exploration through multimodal engagement and adaptive guidance strategies. This evolution continued with the development of socially interactive platforms like Pepper, designed to enhance the visitor experience through storytelling and socially responsive behaviors [27]. However, despite the increasing deployment of such systems, many still rely on prescribed dialogues, limited conversational modules, or human-assisted navigation—factors that constrain their ability to support fluid and meaningful visitor interactions. While significant progress has been made in autonomous navigation, conversational capabilities often remain secondary or disconnected from physical movement. This disconnect can hinder

a robot's perceived social presence, especially in dynamic, real-world settings such as museums. In response, recent research has begun to explore integrated systems that combine navigation autonomy with real-time, context-sensitive dialogue. For example, Del Duchetto and colleagues [16] deployed a Scitos G5 robot at The Collection Museum in the UK, enabling naturalistic, long-term interactions across a wide variety of scenarios. A related deployment by the same group [15] emphasized the robot's ability to provide flexible tour experiences through both verbal communication and touchscreen interfaces. Despite strong operational performance, the study reported a decline in sustained user attention during longer interactions, particularly after the first two minutes of guided tours. This finding supports prior evidence (see Novelty Effect, e.g., [34]) suggesting that initial enthusiasm is insufficient to maintain engagement, and underscores the need for more socially aware behaviors that can adapt to user cues in real time.

In addition to system-level functionality, a growing body of research has examined how individual differences shape the perception and acceptance of museum robots. Fuentes-Moraleda et al. [19] conducted a survey with over 400 Spanish museumgoers, identifying perceived usefulness, empathy, and personal engagement as key factors influencing acceptance. Notably, the study found age to be the most significant demographic predictor: visitors under 30 exhibited markedly higher openness toward robotic guides, viewing them as tools for enhancing personalization and cultural enjoyment. In contrast to this, Webster and Ivanov [40] found that while perceived usefulness and emotional expressivity were critical to acceptance, demographic variables such as age and gender had only a marginal impact.

Further to age, prior encounters with robots also appear to shape how people engage with robotic platforms. Research suggests that repeated or memorable exposure can enhance users' sense of familiarity, trust, and willingness to interact [5, 35]. Recent work by Chan [10] sheds light on the dynamics of incidental human-robot encounters in public spaces, where people engage with a robot without prior knowledge or expectation. Observing a robot requesting assistance outside a university building, the study documented a wide range of spontaneous reactions—from helpful engagement to deliberate avoidance—highlighting the importance of prior expectations and situational context in shaping responses to robots. Although incidental encounters frequently occur in public HRI, particularly in cultural environments like museums, their experiential impact remains insufficiently understood.

To address this gap, our study investigates how visitors perceive a fully autonomous robot functioning as a tour guide in a real museum environment. Unlike earlier systems that treat navigation and dialogue as independent modules, our platform enables integrated verbal and spatial interaction, powered by a large language model. We investigate how visitor perceptions are shaped by individual characteristics (i.e., age, prior experience with robots) and by the nature of their first encounter with the robot (anticipated vs. unexpected). To our knowledge, this is among the first studies to explore how both individual differences and the spontaneity of the initial encounter influence perceptions of a fully autonomous tour guide robot in a real-world museum setting.

3 System Overview

Alter-Ego is a mobile dual-arm robot featuring a humanlike upper-body form factor and variable-stiffness actuators, enabling safe and expressive physical interaction with humans [31]. It is designed for autonomous operation in dynamic environments such as public exhibitions, combining navigation, perception, and natural language interaction capabilities.

3.1 Navigation and Context-Aware Speech

The robot employs a ROS-based navigation stack. An initial map is built using Hector SLAM, while localization is handled in real time via Adaptive Monte Carlo Localization (AMCL). The robot plans safe paths using the Move Base package, which integrates obstacle avoidance for dynamic environments. To enhance engagement, the robot can deliver location-specific utterances as it passes near artworks. For instance, when entering a new zone, it announces relevant content (e.g., "We are now in the 'Sails' area, where you can see...").

3.2 Human Detection and Tour Initiation

Alter-Ego uses the YOLOv10-n object detection model [25] to identify nearby humans based on face size and proximity. When a potential visitor is detected, the robot initiates a greeting, introduces itself, and proposes a guided experience. If the person agrees (via verbal confirmation), the robot starts the tour by calling `go_to()` to navigate to the first exhibit. If the user remains silent or declines, the robot stays in place and resumes scanning. A timeout of 120 seconds without user engagement triggers a system reset.

3.3 Conversational AI and Function Calling

Conversational interaction is managed by the GPT-4o mini model [33], integrated via API with function-calling support. This allows the model to dynamically interpret user input and trigger specific robot actions. The following core functions are supported:

- `go_to(destination)`: Navigates to a selected location based on user input and tour history.
- `end_tour()`: Returns to the initial position and ends the session.

To maintain context while optimizing latency, a dynamic prompt strategy is adopted. The prompt includes:

- (1) **Robot identity and role:** Description of Alter-Ego's capabilities.
- (2) **Current location:** The area where the robot is situated.
- (3) **Tour progression:** Visited and unvisited areas.
- (4) **Exhibit metadata:** Titles, artists, and spatial cues.
- (5) **Conversation history:** Past exchanges to support coherence.

This structure is updated as the robot transitions through the environment, keeping interactions relevant while minimizing prompt length.

3.4 Interaction Flow

The interaction is designed to balance autonomy and user agency. Initially, Alter-Ego invites visitors to join a tour and autonomously

guides them through the first two areas, offering brief contextual explanations. Subsequently, visitors are encouraged to choose where to go next or to end the experience at any time. The system keeps track of progress and adapts its suggestions accordingly. Table 1 summarizes key system behaviors.

Table 1: Overview of Alter-Ego's Behavior in Different Interaction States

Condition	Robot Behavior
Visitor detected	Greets and introduces itself
Visitor agrees to tour	Calls <code>go_to(1st area)</code>
Visitor at 1st area	Calls <code>go_to(2nd area)</code>
Visitor asks for specific area	Calls <code>go_to(desired area)</code>
No interaction for 120s	Calls <code>end_tour()</code>
Visitor ends the tour	Calls <code>end_tour()</code>

4 Data Collection

4.1 Exploratory Field Deployment

A first implementation of the system took place in "Galata Museo del Mare"¹, a Maritime Museum in the city of Genoa (Italy) that hosts a variety of exhibits dedicated to the history of navigation. The museum attracts a diverse audience, including local visitors, tourists, and cruise passengers from different cultural and linguistic backgrounds. The robot was deployed on a floor showcasing maritime artworks, where it autonomously guided participants through selected exhibition areas. Following an initial pilot deployment, we identified several challenges related to the setting and sample. The robot was freely placed in the exhibition space, engaging passersby autonomously without prior recruitment. While this enabled naturalistic interactions, visitors often arrived tired after a long museum visit, reducing engagement. Moreover, the museum's high tourist attendance created a linguistically diverse audience, affecting interaction consistency and data quality. This setting also limited data collection to spontaneous feedback rather than standardized measures.

4.2 Experimental Field Deployment

To address contextual challenges identified and improve data quality, we recruited participants (who all spoke Italian) to take part in an interaction with the robot by visiting the Museum. This allowed for a deeper evaluation and ensured meaningful self-reported responses. In parallel, we restructured the interaction so that each participant followed the robot through at least two exhibition areas, standardizing the exposure time and interaction experience. Finally, we introduced the experimental conditions.

4.2.1 Participants and Experimental Design. Thirty-four Italian adults (mean age = 30.5, SD = 12.6; 17 women, 17 men) were recruited through mailing lists. The study was approved by the local ethics committee, and all participants provided informed consent before taking part. Each participant received compensation for their participation and a museum ticket.

¹<https://en.galatamuseodelmare.it/>

Before the experiment, participants were randomly assigned to one of two experimental conditions:

- **Announced Condition (A)** (n = 17): Participants were told during recruitment that they would be interacting with an autonomous robot guiding them through the exhibition.
- **Surprised Condition (S)** (n = 17): Participants were only told they would be participating in an experiment inside a museum, with no mention of the autonomous robot.

This manipulation was designed to reflect realistic visitor scenarios, where some individuals might be aware of the robot's presence due to prior promotion, while others encounter it unexpectedly (as exemplified in the study by [10]).

To validate the manipulation, participants were explicitly asked orally what they were expecting from the experiment immediately after arrival. The experimenter recorded all responses. The manipulation was effective: all participants in the Announced Condition confirmed they were expecting to see the robot, whereas those in the Surprised Condition reported that they had not anticipated its presence and encountered Alter-Ego unexpectedly.

4.2.2 Interaction with the robot. Upon arrival, participants were welcomed to the museum and introduced to Alter-Ego. They engaged in a semi-structured tour that included two mandatory exhibition areas. They were then free to continue the experience or conclude their visit. The complete guided path comprised seven thematic zones in total. During the visit, participants could interact with Alter-Ego by speaking to it, choosing the areas of the exhibition they wished to explore (e.g., "Let's go to the Sails area!"), and asking the robot questions. These questions ranged from specific curiosities about individual paintings (e.g., "In what year was this painting you're showing me made?") to more general inquiries about the exhibition (e.g., "Which is the most beautiful ship depicted in the pictures?"). Using the system described in Section 3, the robot responded contextually to participants' questions. After the interaction, participants completed the questionnaire described in the next section.

The LLM system guiding the robot successfully engaged with all participants, completing the interaction as intended across sessions. All conversational and navigational behaviors were generated in real time by the autonomous system. No human operator intervened to guide or modify the robot's responses. This ensures that all participant perceptions are based on the robot's actual autonomous performance. Although some natural variability occurred across the 34 interactions, all participants followed a common, semi-structured interaction script with the robot. On average, participants visited 5.69 out of 7 exhibit areas during the interaction (SD = 1.79) and spent 19.20 minutes interacting with the robot (SD = 7.27). While interaction-level analyses and additional details (e.g., turn-taking, timing) are reported in our companion paper [21], here we specifically focus on subjective evaluations and their modulation by framing and user characteristics.

4.2.3 Research Questions. Building on our previous study, which found no direct effect of the experimental condition on visitors' evaluations, we focused on how individual differences may shape responses to a humanoid museum guide. In line with prior research highlighting the moderating role of user characteristics in HRI, we

explored whether *age* and *previous experience with robots* influence the effect of the experimental condition. Specifically, we asked: **RQ1:** Does age influence how the experimental condition affects visitors' perception of the robot? **RQ2:** Does previous experience with robots shape visitors' evaluations in this context?

5 Measurement

To answer our RQs, we assess the participants' impression of the robot Alter-Ego and the evaluation of the interaction with it in a real-life setting. A questionnaire was implemented using items previously used in HRI. We administer the surveys through the open-source tool Sosci Survey². When not specified, the response format was a 7-point Likert scale. When using the term "adapt," it means the items were slightly adjusted from their original version to suit this context. The complete list of items is available in the Additional Materials.

- **Demographics:** gender, age, nationality, degree.
- **Previous experiences with robots:** self-generated multiple choice questions to understand previous experience with robots in general (ranging from 1= "I've never seen one before" to 4= "I work with robots every day");
- **Previous experiences with robot Alter-Ego:** self-generated multiple choice (yes/no) "Have you ever seen Alter-Ego? If yes, where?";
- **Previous Experimental Study with Alter-Ego:** multiple choice (yes/no) "Have you ever taken part in an experiment with Alter-Ego?";
- **Perceived role of the robot:** multiple choice questions about the first-impression of the robot ("Alter-Ego for me is like: a toy; a museum assistant; a mechanical component of the museum; a museum installation) adapted to this context from [2].
- **Human-like appearance** of the robot [17]: 6-item scale on how much they feel the robot was appearing looking human-like.
- **Agency and Experience** [22]: 8-item scale on how much they believed the robot was able to act and feel.
- **Warmth and Competence** [18]: 12-item scale on how much they perceived the robot as competent and friendly.
- **Attitude towards technology** [24] 3-item scale on attitude towards the robot.
- **IOS:** single item on the perception of inclusion of the other in the self towards the robot. [4].
- **Evaluation of Robot's Movements:** 8 items adapted from the scale by [6] evaluating the movements of the robots during the interaction in terms of fluency and adequacy.
- **Evaluation of Robot's Software:** 6 items adapted from the scale by [6], assessing the perceived autonomy and decision-making ability of the robot during navigation in the environment.
- **Intention of Use:** 3 items from [24] to evaluate the participants' willingness to use again the robot in a museum visit.

²<https://www.soscisurvey.de/>

- **Trust in Information:** 5-item scale self-generated for this study to evaluate the trustworthiness, accuracy, and knowledge of Alter-Ego when explaining paintings.
- **Perceived Enjoyment:** 5-items scale from [24] to evaluate the participants' enjoyment during the interaction with Alter-Ego.
- **Perceived Sociability:** 4-items scales from [24]. This scale assesses the perceived likability, understanding, and conversational quality of the robot.
- **Perceived Utility:** 4-items scale adapted from [24] to evaluate how much participants felt Alter-Ego was useful in the museum visit.
- **Interactiveness:** 4-items scale self-generated, to rate how much participants felt the interaction with Alter-Ego was interactive.

To further evaluate participants' general evaluation of the experience, we adapted some slider scales from [3] that allowed participants to respond with a number ranging from 0 (minimum value) to 100 (maximum value).

- **Evaluation of the Robot and the Museum:** We asked participants to express their attitude toward the robot and the Museum on two slider scales, specifying that 0 meant a very unfavorable attitude, 50 a neutral attitude, and 100 a very favorable attitude.
- **Probability of Other Visit:** We asked participants to express the probability of visiting another museum if the robot Alter-Ego was present, specifying that 0 meant "highly unlikely" and 100 meant "highly likely".

Humanlike appearance, Agency and Experience, Warmth and Competence, and IOS scales were administered before and after the interaction; the difference between pre and post evaluations are reported in our previous work [21].

6 Data Analysis

To answer our research questions we have conducted statistical analyses using the open-source statistical software Jamovi³ over the questionnaire results.

6.1 Descriptives

First, we calculated Cronbach's alpha [7] to assess the internal consistency of the questionnaire's multiple item-scales. Mean scores and standard deviations were computed for each experimental condition. We then conducted a series of descriptive analyses to obtain a comprehensive overview of participants' backgrounds and perceptions.

6.2 Moderation Analysis

To investigate potential moderation effects, we conducted a series of General Linear Models (GLMs) using the package GAMLj3 [20] in Jamovi to examine whether the experimental condition interacted with participants' age and prior experience with robots in influencing various dependent variables. Specifically, we tested for interaction effects on measures including perceived warmth, enjoyment of the interaction, overall evaluation of the robot, and

perceived closeness (as measured by the IOS scale). Each GLM included the main effects of condition and the respective covariate (age or prior experience), as well as their interaction term. This approach allowed us to assess whether the relationship between the experimental condition and each dependent variable was moderated by participants' age or their prior experience with robots. No significant moderation effects were found for other dependent variables assessed in the study.

6.3 MANCOVA Analysis

To better interpret the results from the Moderation analysis, we performed an exploratory MANCOVA analysis to get a representation of the relationships among variable collected. We followed the principle of parsimony as recommended by Tabachnick and Fidell [39] to reduce the number of variables included in the model. This approach allowed us to achieve a more interpretable and theoretically meaningful depiction of variable associations while minimizing the risk of overfitting and multicollinearity in models with a relatively small sample size. We conducted a MANCOVA with the experimental condition (Announced vs. Surprised) as the fixed factor and a comprehensive set of dependent variables and covariates. Following a stepwise exclusion approach, we progressively removed covariates that did not significantly affect any of the dependent variables, using a threshold of $p < .05$. Only covariates that significantly impacted at least one DV were retained in the final model. Finally, we checked the direction of the correlation to understand whether the covariates positively or negatively affected the dependent variables. With the method described above, we performed a MANCOVA test in which we tested whether our dependent variables (DV: Agency, Experience, Warmth, Competence, Attitude Towards Technology, Human-like Appearance, Evaluation of Robot Movements, Evaluation of Robot Software, Trust in Information, Interactiveness, Intention of Use, Enjoyment, Perceived Sociability, Perceived Utility, Evaluation of the Robot, Evaluation of the Museum, Probability of visiting another Museum and IOS) were significantly affected by the participants' covariates, (covariates: age, Previous experience with Robots).

7 Results

7.1 Descriptive

A smaller portion of participants had encountered the robot through social media (11.8%) or television (8.8%). A very few mentioned combinations, such as TV and social media (5.9%), or other sources (2.9%). Only one participant (2.9%) reported seeing the robot in the lab setting. All participants declared they had never interacted in an experiment with the robot Alter-Ego before. The majority of participants (67.6%) perceived the robot as an assistant, making it the most frequently selected role. Few viewed the robot as either a component (14.7%) or an installation (14.7%), while only one participant (2.9%) identified the robot as a toy. **The average score for participants' prior experience with the robot was 2.41 (SD = 0.66) on the scale from 1 to 4.**

As reported in Table 2, all scales demonstrated a good level of internal consistency (all $\alpha > 0.70$ ⁴). In the same Table we also

³<https://www.jamovi.org/>

⁴Cronbach's alpha was not calculated for IOS, as it is a single-item measure.

Table 2: Cronbach’s α , mean (for condition), and standard deviation for self-reported scales

Scale (1-7 Likert)	α	Mean A	Mean S
Human-like App.	0.80	1.94 ± 1.21	2.29 ± 0.99
Agency	0.83	3.94 ± 1.57	4.32 ± 1.16
Experience	0.96	2.04 ± 1.44	2.26 ± 1.19
Warmth	0.80	4.75 ± 0.98	5.22 ± 0.69
Competence	0.86	5.21 ± 0.99	5.58 ± 0.69
Attitude Tech.	0.89	5.22 ± 0.93	5.24 ± 0.87
IOS	-	2.76 ± 1.35	3.29 ± 0.92
Eval. Rob. Mov.	0.76	4.35 ± 0.98	4.65 ± 0.81
Eval. Rob. Soft.	0.73	5.26 ± 0.84	5.18 ± 0.88
Trust in Info	0.72	5.60 ± 1.03	5.21 ± 0.67
Interactiveness	0.71	5.21 ± 1.27	5.57 ± 0.83
Intention of Use	0.96	4.39 ± 1.92	4.84 ± 1.54
Enjoyment	0.75	5.44 ± 1.05	5.92 ± 0.68
Perceived Sociability	0.82	4.59 ± 1.31	5.06 ± 0.72
Perceived Utility	0.95	5.26 ± 1.34	5.57 ± 1.14

Table 3: Mean and standard deviation for slider measures

Measure (1-100)	Mean A	Mean S
Evaluation of Alter-Ego	79.90 ± 14.90	81.60 ± 13.50
Evaluation of the Museum	87.60 ± 15.10	83.20 ± 14.70
Probability of Other Visit	66.40 ± 25.40	74.30 ± 19.60

report the mean and standard deviation for each experimental condition. In Table 3, we report the mean and standard deviation of the slider measures. As reported in our previous work, no significant statistical difference has emerged from the Independent Sample T-Tests analysis between the values in the two conditions.

7.2 Moderation

The only significant effect observed across analyses was the **interaction** between the condition (Announced vs Surprised) and covariates (age and prior robot experience); no main effects emerged. Graphs show separate regression lines for Condition A (blue) and Condition S (green), with 95% confidence intervals (shaded) and individual scores (dots).

7.2.1 Interaction Effect: Age x Condition. A significant interaction emerged between condition and age in predicting Evaluation of Alter-Ego, $F(1, 30) = 8.996, p = .005, \eta_p^2 = 0.231$. A simple slopes analysis showed that at low age (−1 SD), participants rated Alter-Ego significantly higher in the Surprised condition ($M_s = 88.22, SE = 4.37$) compared to the Announced condition ($M_a = 72.01, SE = 4.72$), $b = 16.21, p = .017, \eta_p^2 = 0.175$. At high age (+1 SD), the pattern reversed: participants in the Announced condition gave higher ratings ($M_a = 84.18, SE = 3.66$) than those in the Surprised condition ($M_s = 69.43, SE = 6.44$), although this difference only approached significance, $b = -14.75, p = .056, \eta_p^2 = 0.117$ (see Figure 2). No significant difference was found at the mean (M) age level, $b = 0.73, p = .876$.

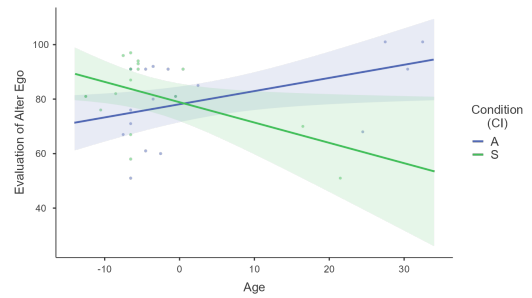


Figure 2: Simple slopes analysis illustrating the interaction between age and condition (Announced vs. Surprised) on the evaluation of the robot. The x-axis represents age (centered at three values: one standard deviation below the mean, the mean, and one standard deviation above the mean). The y-axis indicates enjoyment scores.

Additionally, a significant interaction between condition and age in predicting Enjoyment emerged, $F(1, 30) = 5.53, p = .025, \eta_p^2 = 0.156$. A simple slopes analysis showed that at low age (−1 SD), participants reported significantly higher enjoyment in the Surprised condition ($M_s = 6.08, SE = 0.27$) than in the Announced condition ($M_a = 4.76, SE = 0.29$), $b = 1.32, p = .002, \eta_p^2 = 0.268$. At mean age, the difference between conditions was marginally significant ($b = 0.57, p = .056, \eta_p^2 = 0.116$), while at high age (+1 SD), no significant difference was found ($b = -0.18, p = .694, \eta_p^2 = 0.005$) (see Figure 3).

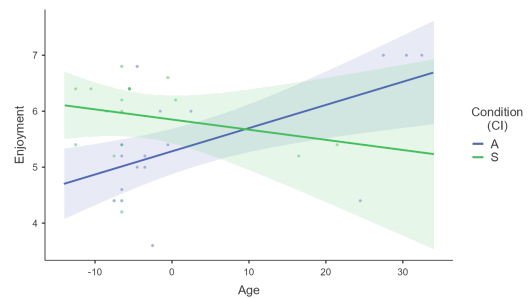


Figure 3: Simple slopes analysis illustrating the interaction between age and condition on enjoyment ratings. The x-axis represents age (centered at three values: one standard deviation below the mean, the mean, and one standard deviation above the mean). The y-axis indicates enjoyment scores.

7.2.2 Interaction Effect: Experience with Robots x Condition. A significant interaction between condition and prior experience with robots emerged for Warmth, $F(1, 30) = 6.03, p = .020, \eta_p^2 = 0.167$. Simple slope analysis revealed that the effect of condition was significant for participants with low experience with robots (−1 SD) ($b = 1.23, SE = 0.43, p = .007$), while it was non-significant for participants with average (M) ($b = 0.48, p = .092$) and higher experience (+1 SD) ($b = -0.26, p = .511$). As shown in Figure 4, participants with low prior experience perceived the robot

as significantly warmer in the Surprised condition ($M_S = 5.49$) compared to the Announced condition ($M_A = 4.27$).

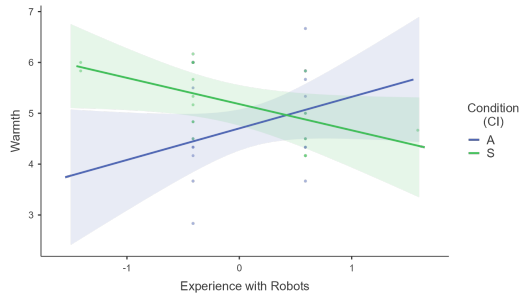


Figure 4: Simple slopes analysis illustrating the interaction between experience with the robot and condition on warmth ratings. The x-axis represents participants' experience with the robot, while the y-axis indicates warmth scores.

Similarly, a significant interaction between condition and prior experience with robots emerged for IOS, $F(1, 30) = 4.21, p = .049, \eta_p^2 = 0.123$. Simple slope analysis revealed that the effect of condition was significant only for participants with low experience with robots (-1 SD) ($b = 1.48, SE = 0.60, p = .020$), while it was non-significant for those with average (M) ($b = 0.60, p = .134$) and higher experience ($+1$ SD) ($b = -0.28, p = .622$). As shown in Figure 5, participants with low prior experience perceived significantly higher closeness with the robot in the Surprised condition ($M_S = 3.49$) compared to the Announced condition ($M_A = 2.01$).

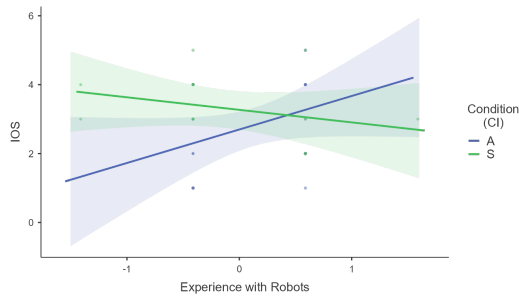


Figure 5: Simple slopes analysis illustrating the interaction between experience with the robot and condition (A vs. S) on IOS scores. The x-axis represents participants' experience with the robot, while the y-axis indicates IOS ratings.

7.3 MANCOVA

The MANCOVA described above revealed that age and prior robot experience influence participants' attitudes toward technology, their experience ratings (i.e., how much they felt the robot was able to feel), and evaluations of robot movements. All other dependent variables and covariates that showed no significant effects were removed from the analysis to ensure model parsimony. The MANCOVA univariate tests indicated that **age significantly predicted attitudes toward technology**, $F(1, 30) = 6.12, p = 0.019$,

while **prior robot experience significantly predicted both the "Robot Experience" rating** (i.e., the capacity of the robot to feel), $F(1, 30) = 4.95, p = 0.034$, and the **evaluation of robot movements**, $F(1, 30) = 6.32, p = 0.018$. The experimental condition did not significantly affect any of the dependent variables (all $p > 0.29$). Regarding the direction of these effects, age was positively correlated with attitudes toward technology ($r = 0.384, p = 0.025$), indicating that older participants tend to hold more positive attitudes toward technology. Moreover, prior robot experience was negatively associated both with "Robot Experience" ratings ($r = -0.349, p = 0.043$) and with the evaluation of robot movements ($r = -0.356, p = 0.039$). This suggests that participants with greater prior experience in robotics tend to rate the robot as less able to perceive and feel, and to be critical when evaluating specific robot behaviors or movements. In Table 4 we summarize the results of the MANCOVA analysis.

Table 4: MANCOVA results: effects of age and prior robot experience

Dependent Variable	Predictor	F (df)	p	Direction (r)
Attitude toward technology	Age	6.12 (1,30)	0.019	+0.384
Robot experience evaluation	Prior robot exp.	4.95 (1,30)	0.034	-0.349
Robot movement evaluation	Prior robot exp.	6.32 (1,30)	0.018	-0.356

8 Discussion

This study examined how individual characteristics influence the impact of the experimental condition on the evaluation of the robot. Moderation analyses revealed that **age** and **prior experience with robots** significantly shaped participants' perceptions and moderated the effect of the condition (*Announced* vs. *Surprised*). These findings highlight the important role of individual differences in how people interpret and evaluate robotic behavior. The MANCOVA further confirmed the influence of Age and Experience on the evaluations of Alter-Ego.

8.1 Age Influence

The results showed that age significantly moderated the relationship between the experimental condition and both the evaluation of the Alter-Ego and the enjoyment of the interaction, providing an answer to **RQ1**. Younger participants evaluated the robot more positively and reported greater enjoyment in the **Surprised** condition, where the robot's behavior was not introduced in advance. In contrast, older participants showed a marginally better evaluation of the robot in the **Announced** condition. These findings suggest an age-related difference in user preferences: younger participants responded more positively when the interaction with the robot was unanticipated, whereas older participants showed a preference for scheduled interactions. Furthermore, results from the MANCOVA analysis seem to confirm a lesser enthusiasm from the younger sample: age significantly and positively predicted attitudes toward technology, meaning that older participants in our

sample reported more positive attitudes towards the robot than younger ones. These results differ from previous HRI research identifying age as a key factor influencing attitudes and behaviors toward robots (e.g., [29]), including in museum contexts [38, 40]. Specifically, [19] found that younger museum visitors were more likely to accept robots, an opposite trend to what we observed in our study. Younger users preferred the interaction when the robot appeared unexpectedly, reporting greater enjoyment in this condition. They evaluated the robot more negatively when its presence was announced and showed a less positive attitude towards the technology compared to older users. A possible explanation for this pattern may come from Socioemotional Selectivity Theory [8], which suggests that motivational priorities shift across the lifespan. Older adults tend to prioritize emotionally meaningful, predictable experiences, while younger individuals are more oriented toward novelty and exploratory behaviors. In the context of our study, the structured interaction with the announced robot may have aligned more closely with older visitors' motivational goals, enhancing their evaluations. Conversely, the unexpected presence of the robot may have been more stimulating, and therefore enjoyable, for younger participants.

8.2 Previous Experience with Robots Influence

Our results show that prior robot experience moderated participants' perceptions of the robot's warmth and their reported social closeness (IOS scale), addressing **RQ2**. Participants with less experience perceived the robot as warmer and felt more socially connected when the interaction was announced, suggesting that contextual framing benefits less experienced users. In contrast, more experienced participants' evaluations remained stable, likely due to existing expectations or familiarity.

MANCOVA analyses further revealed that prior experience influenced **mind attribution** and **movement quality** evaluations. Interestingly, more experienced participants made more critical assessments, attributing fewer experiential capacities and rating movements more critically.

Prior experiences in robotics have been shown to influence participants' behavior and evaluation of robots [12] [35]. For example, Saunderson and Nejat [35] found that prior exposure to a robot enhanced its perceived trustworthiness and persuasive impact. However, previous experience with robots does not always lead to positive outcomes. Longitudinal research [28] revealed that greater familiarity influenced enjoyment of the interaction in a non-linear pattern—peaking early, dipping, and then rising again over time. These results are consistent with our findings: visitors with greater prior experience, either from participating in previous robot experiments or having broader general knowledge about robots (1) were not influenced by the experimental condition when evaluating the robot, and (2) tended to give more negative evaluations regarding the robot's mental capacities (e.g., its ability to feel) and movements. It is plausible that experienced users approach the interaction with more grounded expectations, relying less on contextual cues, while novice users are more susceptible to framing effects. Another plausible explanation comes from the "Novelty Effect" [37], whereby repeated exposure reduces curiosity and engagement—an effect previously documented in HRI literature [14, 30]. Our results suggest

that experienced visitors are less influenced by contextual framing, while novices show greater sensitivity and enthusiasm. Prior experience appears to shape how people interpret encounters with autonomous agents.

9 Conclusions

In our previous study, announcing the robot's presence did not prove to significantly influence participants' evaluations. Current findings refine this view: whether or not the robot's announcement influences participants' responses appears to depend on factors such as **age** and **prior familiarity with robots**. Individual differences moderate how interaction framing affects social robot perception. Our findings indicate possible age-related preferences regarding how the robot is introduced, which diverge from patterns reported in previous studies. Prior robot experience also shapes evaluations, with more experienced users considering contextual factors. Both findings warrant further investigation, indicating that user background and expectations are crucial for optimizing interactions.

Some limitations should be noted. First, the relatively small sample size (N=34) may limit generalizability and power. Despite this, the moderation analyses revealed interpretable interaction patterns that suggest preliminary trends worthy of further investigation. Moreover, relying solely on self-reports can introduce biases such as social desirability [23]; future studies should also combine behavioral or physiological measures for a broader understanding. Finally, future research should also consider other personal traits, such as gender and anxiety, as they may further influence how users perceive and interact with robots.

Despite these limitations, our findings offer valuable insight into how autonomous robots are received in everyday cultural settings. Importantly, by conducting our study **in the wild**, we contribute to bridging the gap between controlled laboratory research and real-world HRI, capturing how people actually engage with robots in complex, naturalistic settings.

Beyond its empirical contributions, this research highlights the value of a **multidisciplinary** approach to HRI. By combining robotic implementation with social psychology methods, it offers a more comprehensive and integrative view on emerging social meaning in interactions with autonomous systems, as emphasized in prior HRI literature [13] [1].

Ultimately, this research underscores the need to design and evaluate autonomous robots within the complex interplay of social, psychological, and contextual factors shaping user perception. These insights support more adaptive deployment strategies in public settings, where personalization and expectation management are key to successful integration.

10 Acknowledgements

This work was supported by: "RAISE – Robotics and AI for Socioeconomic Empowerment" (GRANT NUMBER) and by the European Union - NextGenerationEU (GRANT NUMBER). Thanks to Mattia Poggiani, Cristiano Petrocelli, and Manuel Barbarossa for technical support, and especially to Giovanni Rosato and Eleonora Sguerri for technical assistance and help during data collection.

References

- [1] Patrícia Alves-Oliveira, Dennis Küster, Arvid Kappas, Ana Paiva, et al. 2016. Psychological Science in HRI: Striving for a More Integrated Field of Research. In *AAAI Fall Symposia*.
- [2] Patrícia Alves-Oliveira, Pedro Sequeira, and Ana Paiva. 2016. The role that an educational robot plays. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 817–822.
- [3] Duane F Alwin. 1997. Feeling thermometers versus 7-point scales: Which are better? *Sociological Methods & Research* 25, 3 (1997), 318–340.
- [4] Arthur Aron, Elaine N Aron, and Danny Smollan. 1992. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology* 63, 4 (1992), 596.
- [5] Ritta Baddoura, Gentiane Venture, and Ryo Matsukata. 2012. The familiar as a key-concept in regulating the social and affective dimensions of HRI. In *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)*. IEEE, 234–241.
- [6] Jasmin Bernotat, Lorenzo Landolfi, Dario Pasquali, Alice Nardelli, and Francesco Rea. 2023. Remember me-user-centered implementation of working memory architectures on an industrial robot. *Frontiers in Robotics and AI* 10 (2023), 1257690.
- [7] Mohamad Adam Bujang, Evi Diana Omar, and Nur Akmal Baharum. 2018. A review on sample size determination for Cronbach's alpha test: a simple guide for researchers. *The Malaysian journal of medical sciences: MJMS* 25, 6 (2018), 85.
- [8] Laura L Carstensen. 2021. Socioemotional selectivity theory: The role of perceived endings in human motivation. *The Gerontologist* 61, 8 (2021), 1188–1196.
- [9] Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. 2012. Research in the wild: understanding' in the wild' approaches to design and development. In *Proceedings of the designing interactive systems conference*. 795–796.
- [10] Yao-Cheng Chan. 2025. A Field Observation of Incidental Human-Robot Encounters in Public. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1265–1268.
- [11] Antonio Chella, Marilia Liotta, and Irene Macaluso. 2007. CiceRobot: a cognitive robot for interactive museum tours. *Industrial Robot: An International Journal* 34, 6 (2007), 503–511.
- [12] Francesca Ciardo, Davide Ghiglinò, Cecilia Roselli, and Agnieszka Wykowska. 2020. The effect of individual differences and repetitive interactions on explicit and implicit attitudes towards robots. In *International conference on social robotics*. Springer, 466–477.
- [13] Emily C Collins. 2019. Drawing parallels in human–other interactions: a trans-disciplinary approach to developing human–robot interaction methodologies. *Philosophical Transactions of the Royal Society B* 374, 1771 (2019), 20180433.
- [14] Kerstin Dautenhahn. 2007. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical transactions of the royal society B: Biological sciences* 362, 1480 (2007), 679–704.
- [15] Francesco Del Duchetto, Paul Baxter, and Marc Hanheide. 2019. Lindsey the tour guide robot-usage patterns in a museum long-term deployment. In *2019 28th IEEE international conference on robot and human interactive communication (RO-MAN)*. IEEE, 1–8.
- [16] Francesco Del Duchetto, Paul Baxter, and Marc Hanheide. 2020. Are you still with me? Continuous engagement assessment from a robot's point of view. *Frontiers in Robotics and AI* 7 (2020), 116.
- [17] Francesco Ferrari, Maria Paola Paladino, and Jolanda Jetten. 2016. Blurring human–machine distinctions: Anthropomorphic appearance in social robots as a threat to human distinctiveness. *International Journal of Social Robotics* 8, 2 (2016), 287–302.
- [18] Susan T Fiske, Amy JC Cuddy, and Peter Glick. 2007. Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences* 11, 2 (2007), 77–83.
- [19] Laura Fuentes-Moraleda, Carmen Lafuente-Ibanez, Natalia Fernandez Alvarez, and Teresa Villace-Moliner. 2022. Willingness to accept social robots in museums: an exploratory factor analysis according to visitor profile. *Library Hi Tech* 40, 4 (2022), 894–913.
- [20] Marcello Gallucci. 2019. GAMLj: General Analyses for Linear Models. <https://gamlj.github.io/>. [jamovi module].
- [21] Luca Garello, Francesca Cocchella, Alessandra Sciutti, Manuel Catalano, and Francesco Rea. 2025. Next-Gen Museum Guides: Autonomous Navigation and Visitor Interaction with an Agentic Robot. *arXiv preprint arXiv:2507.12273* (2025).
- [22] Heather M Gray, Kurt Gray, and Daniel M Wegner. 2007. Dimensions of mind perception. *Science* 315, 5812 (2007), 619–619.
- [23] Pamela Grimm. 2010. Social desirability bias. *Wiley international encyclopedia of marketing* (2010).
- [24] Marcel Heerink, Ben Krose, Vanessa Evers, and Bob Wielinga. 2009. Measuring acceptance of an assistive social robot: a suggested toolkit. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 528–533.
- [25] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. 2023. *YOLO by Ultralytics*. <https://github.com/ultralytics/ultralytics>
- [26] Malte Jung and Pamela Hinds. 2018. Robots in the wild: A time for more robust theories of human-robot interaction. 5 pages.
- [27] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. 2004. Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction* 19, 1-2 (2004), 61–84.
- [28] Aelee Kim, Jooyun Han, Younbo Jung, and Kwanmin Lee. 2013. The effects of familiarity and robot gesture on user acceptance of information. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 159–160.
- [29] Melissa Kont and Maryam Alimardani. 2020. Engagement and mind perception within human-robot interaction: A comparison between elderly and young adults. In *Social Robotics: 12th International Conference, ICSR 2020, Golden, CO, USA, November 14–18, 2020, Proceedings 12*. Springer, 344–356.
- [30] Iolanda Leite, Carlos Martinho, and Ana Paiva. 2013. Social robots for long-term interaction: a survey. *International Journal of Social Robotics* 5 (2013), 291–308.
- [31] Gianluca Lentini, Alessandro Settimi, Danilo Caporale, Manolo Garabini, Giorgio Grioli, Lucia Pallottino, Manuel G Catalano, and Antonio Bicchi. 2019. Alter-ego: a mobile robot with a functionally anthropomorphic upper body designed for physical interaction. *IEEE Robotics & Automation Magazine* 26, 4 (2019), 94–107.
- [32] Marco Matarese, Francesca Cocchella, Francesco Rea, and Alessandra Sciutti. 2023. Natural Born Explainees: how users' personality traits shape the human-robot interaction with explainable robots. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1786–1793.
- [33] OpenAI. 2024. *GPT-4o mini: advancing cost-efficient intelligence*. <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>
- [34] Merle Reimann, Jesper van de Graaf, Nina van Gulik, Stephanie Van De Sanden, Tibert Verhagen, and Koen Hindriks. 2023. Social robots in the wild and the novelty effect. In *International Conference on Social Robotics*. Springer, 38–48.
- [35] Shane Saunderson and Goldie Nejat. 2021. Robots asking for favors: The effects of directness and familiarity on persuasive HRI. *IEEE Robotics and Automation Letters* 6, 2 (2021), 1793–1800.
- [36] Roland Siegwart, Kai O Arras, Samir Bouabdallah, Daniel Burnier, Gilles Froidevaux, Xavier Greppin, Björn Jensen, Antoine Lorotte, Laetitia Mayor, Mathieu Meisser, et al. 2003. Robox at Expo. 02: A large-scale installation of personal robots. *Robotics and Autonomous Systems* 42, 3-4 (2003), 203–222.
- [37] Catharina Vesterager Smedegaard. 2019. Reframing the role of novelty within social HRI: from noise to information. In *2019 14th acm/ieee international conference on human-robot interaction (hri)*. IEEE, 411–420.
- [38] Richard T Sweeney. 2005. Reinventing library buildings and services for the millennial generation. *Library Leadership & Management* 19, 4 (2005), 165–176.
- [39] BG Tabachnik and SL Fidell. 2007. Multivariate normality. *Using multivariate statistics* 6 (2007), 253.
- [40] Craig Webster and Stanislav Ivanov. 2022. Public perceptions of the appropriateness of robots in museums and galleries. *Journal of Smart Tourism* 2, 1 (2022), 33–39.