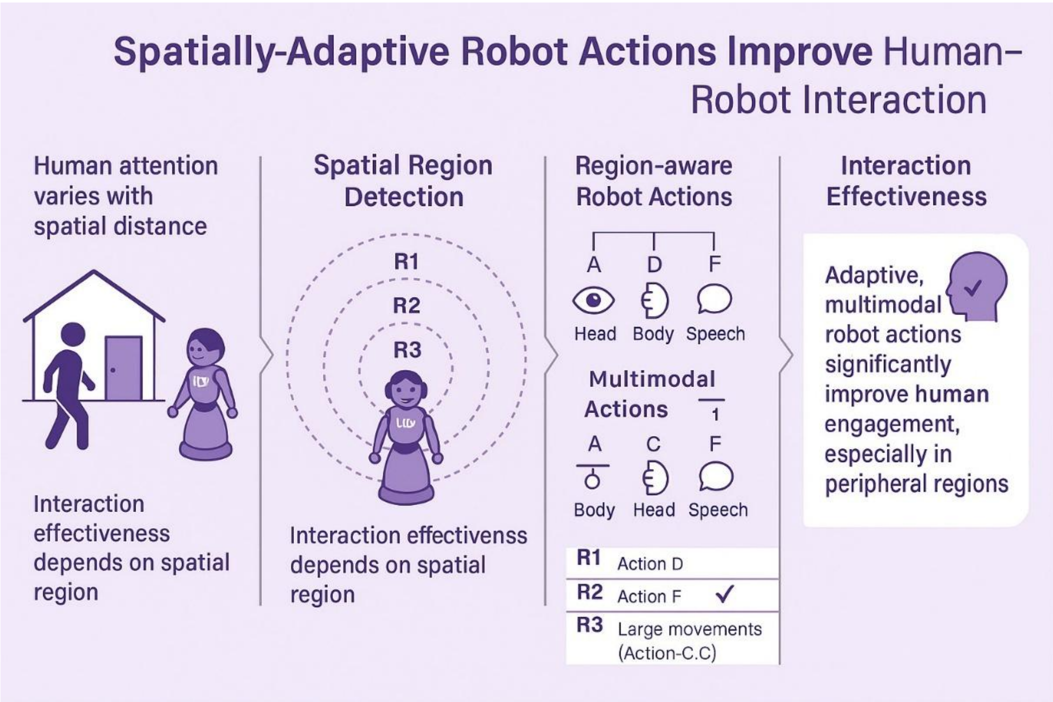
Optimizing Human-Robot Interaction: An Analysis of Robot Action Effectiveness in Different Spatial Regions

D. A. A. Deepal[1,2,*], N. H. Wanigasingha[1], M. K. A. Ariyaratne[1] and P. R. S. De Silva[1]

[1]*Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Sri Lanka,*
[2]*Faculty of Graduate Studies, University of Sri Jayewardenepura, Gangodawila, Sri Lanka.*

Date Received: 14-06-2025 Date Accepted: 25-12-2025



Abstract:

This study investigates the effectiveness of various robot actions in facilitating human-robot interactions across different spatial regions. The research focuses on a social robot, 'Lily,' designed to engage visitors at a smart home entrance. By employing human attention shift and interaction success analysis, we examine the impact of specific robot actions, including head movements and verbal cues, in different proximity zones (central, near peripheral, and peripheral regions). Our findings suggest that dynamic, multimodal actions, particularly those involving body rotation and verbal cues (Action F), are most effective in peripheral regions, where human engagement is typically lower. In contrast, simpler actions, such as head movements with verbal behaviors (Action D), proved more successful in central regions with closer human proximity. The study highlights the importance of adapting robot behaviors based on spatial positioning to optimize engagement and communication effectiveness. These insights offer valuable guidance for designing robots that can dynamically adjust their interactions based on real-time spatial context, improving Human-Robot Interaction (HRI) in public and domestic environments.

Keywords: *Human–Robot Interaction, Social Robotics, Spatial Proxemics, Multimodal Interaction, Attention Shift Analysis, Adaptive Robot Behavior*

**Correspondence:* deepal@sjp.ac.lk

1. Introduction

A social robot consists of two main features which are called robot and social interface. The social interface encompasses all the designed features that contribute to the robot's social qualities and capabilities, including physical appearance, communication abilities, emotional expression, social awareness, and adaptive behaviors, which are intentionally implemented to enable the robot to engage in appropriate and meaningful social interactions within human contexts (Hegel et al., 2009). A social robot must exhibit contextually appropriate behaviors and possess a distinct form that enhances user interaction (Breazeal et al., 2016). The social interface consists of social function, social forms, and social context. These parameters can be used as designation guidelines when designing a social robot. Forms are the elements that can be sent by social signs and signals physically. These forms play a crucial role in enhancing human-robot communication by conveying non-verbal cues. For example, the design of the face of the social robot is very important because the face of the social robot can be used to show some non-verbal signs and signals. Artificial social behavior (emotions), joint attention mechanism and module of speech recognition are part of the social functions. These components facilitate dynamic and adaptive social interactions between humans and robots. Features of social context conclude form and function. A robot's application defines its operational context and influences its functional design (Hegel et al., 2009).

Householders frequently encounter interruptions from unexpected visitors, which can disrupt daily routines and privacy. Whether the visitor is essential (e.g., close relatives) or non-essential (e.g., solicitors), the householder may be occupied with urgent tasks and unable to attend the door immediately. Consequently, there is a need for an intelligent gatekeeping system that can autonomously manage these interactions based on the visitor's importance.

To mitigate such issues, the main intention of the study is to develop a sociable robot which could be developed to deal with important guests and entertain them until the householder is free. Additionally, the robot will be designed to handle unwanted guests using predefined conversational strategies. This robotic system will furthermore lead to strengthening the security aspects of the household by being cautious of the guests and identifying who is important.

While social robotics has advanced significantly, there is a distinct lack of research applying social robotic principles to doorbell systems, particularly those operating in the Sinhala language. Furthermore, few studies have integrated dynamic proxemic behavior adjustment specifically for a gatekeeping robot in a domestic context. This study addresses this dual gap by introducing a Sinhala-speaking robot that adapts its actions based on the visitor's spatial region.

2. Related Work

Face detection, recognition, and natural language processing (NLP) constitute the fundamental technological triad enabling personalized human-robot interaction. These systems allow robots to direct attention, identify individuals, and engage in conversational dialogue, thereby enhancing responsiveness and facilitating meaningful social connections (Avila and Bailey, 2015; Datta and Vijay, 2010; Yoshiike et al., 2010).

Practical implementations showcase these technologies in diverse applications. The Sociable Trash Box (STB) engages children in environmental cleanup through vocal cues and expressive body movements, encouraging interpretation and assistance (Yoshiike et al., 2010). Neel demonstrates integrated functionality with autonomous navigation, touchscreen interaction, and social networking capabilities, showing strong user engagement in field tests (Datta and Vijay, 2010). For domestic settings, Jibo serves as a family assistant employing voice/face recognition and NLP to perform tasks and tell stories (Avila and Bailey, 2015), while Mykie acts as a kitchen assistant with recipe projection and Aido functions as a multi-purpose smart home robot for childcare and scheduling

*Correspondence: deepal@sjp.ac.lk

(https://interestingengineering.com/innovation/15-small-robots-that-will-invade-your-home-sooner-than-you-think).

Human attention operates through complex cognitive and behavioral filters that selectively process verbal and non-verbal cues in social settings (Carraro et al., 2017). This multi-stage process involves perception, evaluation, and interpretation to effectively filter stimuli (Capozzi and Ristic, 2018). For successful human-robot interaction, robots must first capture attention before initiating communication, a process termed attention modulation involving deliberate redirection of human focus (Hoque et al., 2012a).

Significant challenges exist in attention modulation, particularly the uncanny valley effect, where highly human-like robotic features induce discomfort and reduce engagement (Wang et al., 2015). Facial proportions crucially influence perceived familiarity and approachability (DiSalvo et al., 2002), while physical posture and orientation significantly impact social acceptance levels (Kwak, 2014). Eye contact management through both overt and covert mechanisms reinforces the importance of gaze in social interactions (Iwasaki et al 2022; Gobel and Giesbrecht, 2020).

Spatial arrangement profoundly influences attentional responses, with closer proximity facilitating greater focus and interaction (Frijns et al., 2023; Hall et al., 1968). Humans naturally respond to motion-based stimuli, making animated gestures effective for attention capture, though obstructed visual fields can impair non-verbal cue efficacy (Das et al., 2013). In such scenarios, verbal engagement serves as an alternative attention-directing mechanism (Hoque et al., 2014).

Visual Focus of Attention (VFOA) plays a fundamental role in engagement, with robots optimizing interaction timing by tracking gaze direction and head orientation (Das et al., 2013). Individuals periodically shift visual focus based on cognitive load and task complexity (Masse et al., 2018), requiring dynamic adjustment capability from robotic systems. While eye contact remains crucial for communication, direct gaze alignment alone may be insufficient for sustained attention without complementary behaviors like subtle head tilts and blinking (Hayward et al., 2017; Hoque et al., 2012a). Head orientation tracking often provides a more reliable attentiveness indicator than direct eye-tracking (Hoque et al., 2012a), with future HRI advancements needing to refine natural communication techniques while considering personal space and societal acceptance factors (Glas et al., 2015; Hoque et al., 2012b).

Contemporary smart home technologies, including doorbells with motion sensing and video recording, enable remote household management but present limitations requiring active user involvement for access verification. Unanswered calls may expose household absence, creating security concerns. These systems necessitate physical user intervention for operation.

To address these limitations, we implemented Lily, a Sinhala-speaking sociable robot doorbell designed to autonomously recognize guests, engage in natural interactions, and politely refuse unwanted visitors. This research responds to the identified need for further investigation into attentional responses in dynamic, real-world environments (Das et al., 2013), advancing toward more intuitive and secure human-robot interaction systems.

3. Design of the Robot

This study presents a sociable robot doorbell implementing posture and positional modalities based on Hall's proxemics theory and Integrated Attention Control (Hoque et al., 2012a). The minimalist design optimizes efficiency, usability, and maintenance while providing aesthetic appeal through smooth curved edges and a modern purple-white color scheme.

3.1 Minimalist Design

Minimalist design for a robot optimizes efficiency, promotes ease of use, simplifies maintenance and repair, facilitates versatility and adaptability, and provides an appealing aesthetic. These

*Correspondence: deepal@sjp.ac.lk

advantages make a minimalist design highly desirable for a wide range of practical applications. Therefore, in this research, we used a minimalist design for our robot.

The robot comprises three modular components: Head (165mm), Upper Body (310mm), and Lower Body. The Head features full rotational capability and houses an LCD (50mm×76mm window), 90° webcam (28.5mm mount), and an internal speaker. The Upper Body's rotating segment (185mm) contains an Intel RealSense sensor for distance estimation and spatial classification (intimate, personal, social, public spaces). The Lower Body integrates rear-panel connectivity (HDMI, USB, power) and volume control. The complete assembly stands 48cm tall with a 12.5cm diameter (Figure 1).

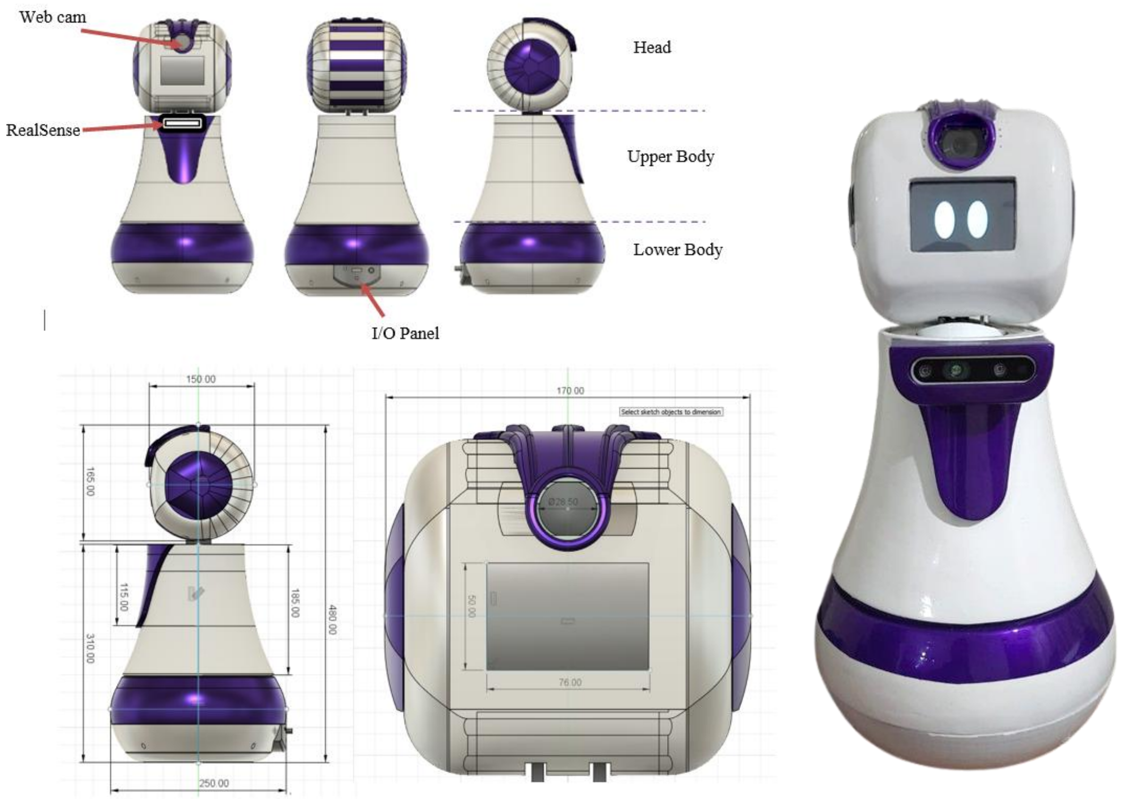


Figure 1. Exterior design of the robot (3D-Printing design)

Dynamic eye expressions (Natural, Happy, Angry, Afraid, Sad, Surprise, Thinking, Listening) were implemented using HTML/CSS/JavaScript (Figure 2), featuring natural blinking and smooth state transitions (Figure 3) to enhance emotional communication.

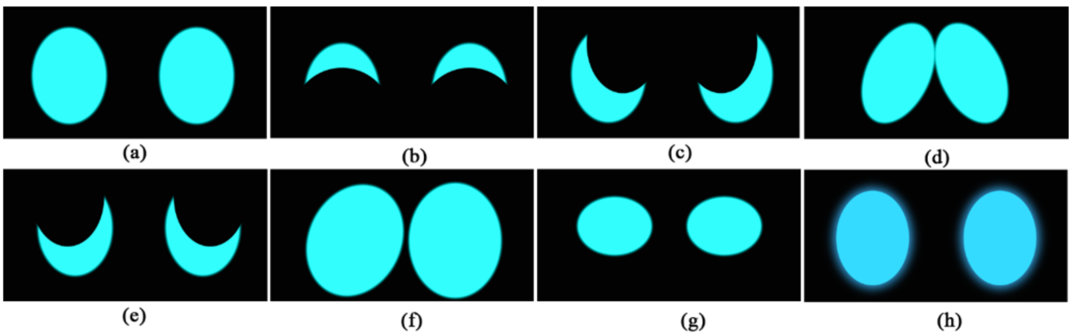


Figure 2. Eye expressions: (a). Natural (b). Happy (c). Angry (d). Afraid (e). Sad (f). Surprise (g). Thinking (h). Listening

*Correspondence: deepal@sjp.ac.lk

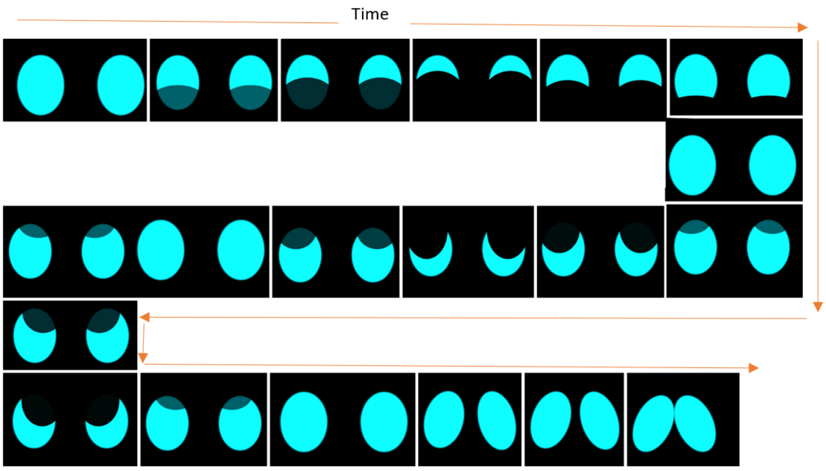


Figure 3. Eye expression: Natural ->happy-> Sad -> Angry -> Surprise

The Upper Body (310 mm) features a rotating upper segment (185 mm) housing an Intel RealSense camera, while the Lower Body provides rear-panel connectivity (HDMI, USB, power) and volume control. The complete assembly stands 48 cm tall with a 12.5 cm diameter, featuring a modern purple and white color scheme that complements its sleek, compact design.

To optimize printing time and structural integrity, the robot was segmented into smaller components. Specific parts were adapted to house the webcam, speaker, and LCD while preserving the original design. Assembly using nuts, bolts, and adhesive ensured a stable and durable final structure, as illustrated in Figure 4.

Figure 4. Sliced parts of the robot

*Correspondence: deepal@sjp.ac.lk

3.2 *Hardware Design and Implementation*

The robot's functionality integrates multiple components: three Dynamixel servo motors, a 3.5-inch LCD display, speaker, Rapoo C260 webcam, Intel® RealSense™ Depth Camera D456, USB 3.0 hub, HDMI joint socket, 7.1 channel USB sound card, LM386 Audio Amplifier Module with volume controller, and Dynamixel servo USB connector.

Motor configuration enables comprehensive movement: one motor in the Head facilitates vertical movement (up/down), a second in the Upper Body enables horizontal head rotation (left/right), and a third in the Lower Body controls Upper Body horizontal rotation (Figure 5). All motors connect via Dynamixel cables to a USB connector integrated with the hub.

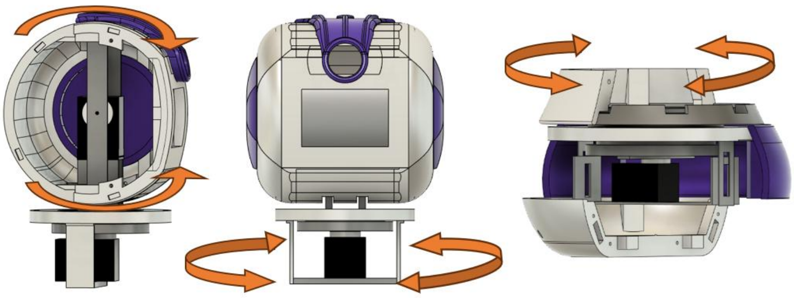


Figure 5. Head and Body Rotation

The Intel RealSense camera, mounted on the Lower Body, serves dual purposes: detecting individuals and estimating distances to classify interaction zones (intimate, personal, social, or public space), while actively tracking guest movements with coordinated head adjustments to maintain eye contact.

Power distribution uses a 12V AC/DC adapter supplying motors and volume controller. The LCD connects via HDMI to the I/O panel and receives power through USB. Audio routing links the sound card adapter to the volume controller, which connects to the internal speaker. The I/O panel consolidates power input, HDMI, USB, and volume control.

A central USB hub in the Lower Body connects the Dynamixel connector, webcam, sound card adapter, and LCD display, while providing PC connectivity for integrated control of motors, audio, and vision systems (Figure 6). This configuration ensures seamless power distribution, audio functionality, and centralized control.

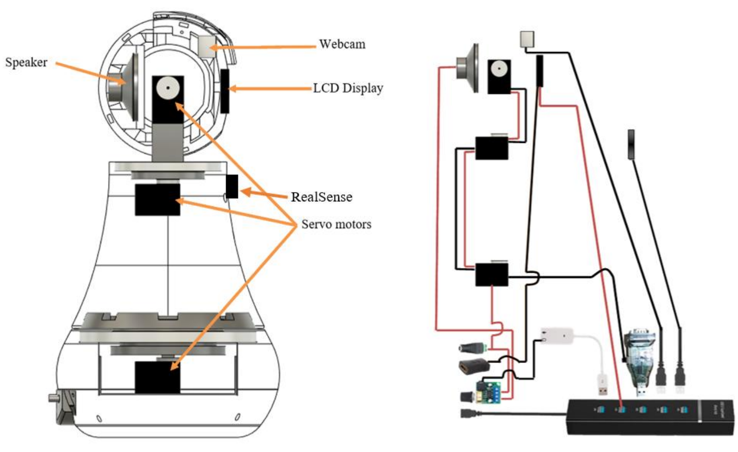


Figure 6. Hardware design

*Correspondence: deepal@sjp.ac.lk

3.3 Software Implementation and Technologies

The system utilizes Intel RealSense SDK 2.0 for human detection and distance measurement, classifying user proximity into Hall's spatial zones (Hall et al., 1968). A YOLOv8 model integrated with OpenVINO (https://docs.openvino.ai/latest/home.html) enables robust person detection. When individuals enter social space, the system initiates interaction.

Facial analysis employs face-api.js (https://justadudewhohacks.github.io/face-api.js/docs/index.html) with TensorFlow.js for detection, recognition, landmark detection, expression, age, and gender estimation. A Single Shot Multibox Detector (SSD) based on MobileNetV1 generates 128-dimensional face descriptors for known/unknown person identification.

Upon face detection, the system extracts comprehensive profile data (identity, demographics, expressions, coordinates) and activates voice/language modules for social space entrants. Motor control via the Dynamixel SDK adjusts head/body orientation to maintain focus based on user coordinates.

Communication leverages a custom SNLP (Sinhala Natural Language Processing) Toolkit built with Python, Node.js, C#, MySQL, and MongoDB. This platform supports Sinhala Named Entity Recognition model training and dialogue flow customization, using similarity scoring for response generation through API integration.

3.4 Behavior Sequencing of the Robot

Building upon Hoque et al.'s (Hoque et al., 2012a) classification of positional relations in human-robot interaction, this study implements five interactive regions based on the guest's position relative to the robot's field of view (Figure 7). The system specifically addresses central field of view (CFOV) and near peripheral field of view (NPFOV) scenarios, divided into distinct interactive regions (Table 1).

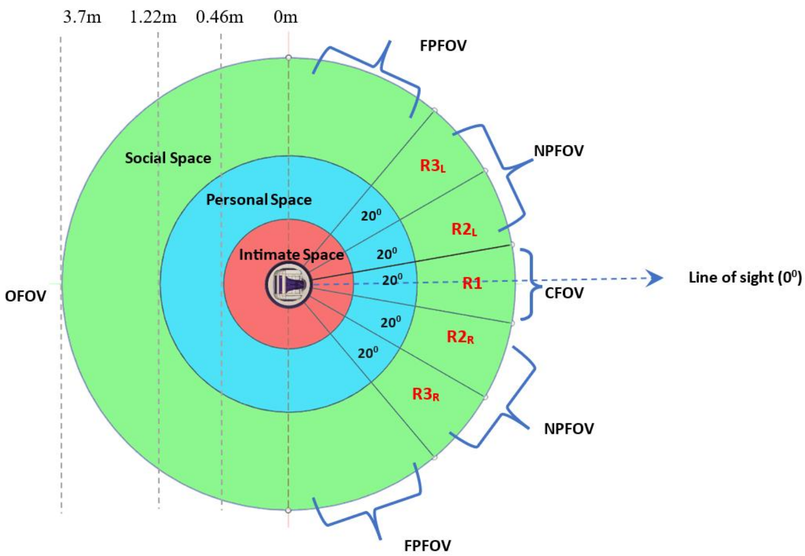


Figure 7. Interactive regions, interpersonal distances and viewing situations of the robot

Six predefined action groups (Table 2), employ escalating combinations of head movements, body rotations, and verbal behaviors to capture attention. When detecting a guest, the robot executes region-specific actions from these predefined groups.

*Correspondence: deepal@sjp.ac.lk

Table 1. Interactive regions and viewing situations of the robot

Interactive regions	positional relations	Angle	Description
R1	CFOV	20^0	Central field of view
R2L	NPFOV	20^0	Left near the peripheral field near CFOV
R2R	NPFOV	20^0	Right near the peripheral field near CFOV
R3L	NPFOV	20^0	Left near the peripheral field near OFOV
R3R	NPFOV	20^0	Right near the peripheral field near OFOV

Table 2. Predefined Action Groups

Action Group	Predefined Action
Action A	The robot moves its head to focus on the center of the detected person's face, turning it up, down, left, or right as needed.
Action B	The robot's body rotates toward the person's direction, then it moves its head to focus on the center of the detected person's face, turning it up, down, left, or right.
Action C	The robot partially rotates its body, then rotates its head toward the person's direction, and finally moves its head to focus on the center of the detected person's face.
Action D	Combination of Action A and verbal behaviors, where the robot speaks to the user.
Action E	Combination of Action B and verbal behaviors, where the robot speaks to the user.
Action F	Combination of Action C and verbal behaviors, where the robot speaks to the user.

3.5 Robot's Action Execution Process

The interactive doorbell 'Lily' continuously monitors its environment using Intel RealSense technology. When a person is detected, the system estimates their position and interaction region, calculating distance to determine engagement needs. If the individual is within the Social Space, a person-detected flag is activated, triggering facial recognition to identify if the guest is known or unknown. Age and gender are also estimated to provide contextual cues.

Angular calculations help optimize the robot's head and body positioning, which motor controllers then adjust based on the guest's location to ensure effective interaction. Engagement begins with the activation of language modules for voice responses and speech recognition.

If the RealSense camera fails to detect the person but the webcam identifies a face in the R3 region, the system compensates by directing motor controllers to rotate the head and body toward the guest. This adaptability allows the robot to overcome RealSense's limited horizontal viewing angle, accurately capturing distance and deciding whether to engage.

When the guest leaves, the robot smoothly returns to its initial position, ready to engage the next visitor efficiently.

*Correspondence: deepal@sjp.ac.lk

4. Experimental Protocol

The experiment aimed to identify optimal robot behaviors for attention capture across spatial regions. The robot was positioned in an open area between Rooms A and B, a natural pathway connecting corridors and staircases (Figure 8). We observed how the robot attracted attention and initiated interactions, recording engagement success across interactive regions. Participants provided qualitative feedback through post-interaction questionnaires to capture user experience perceptions.

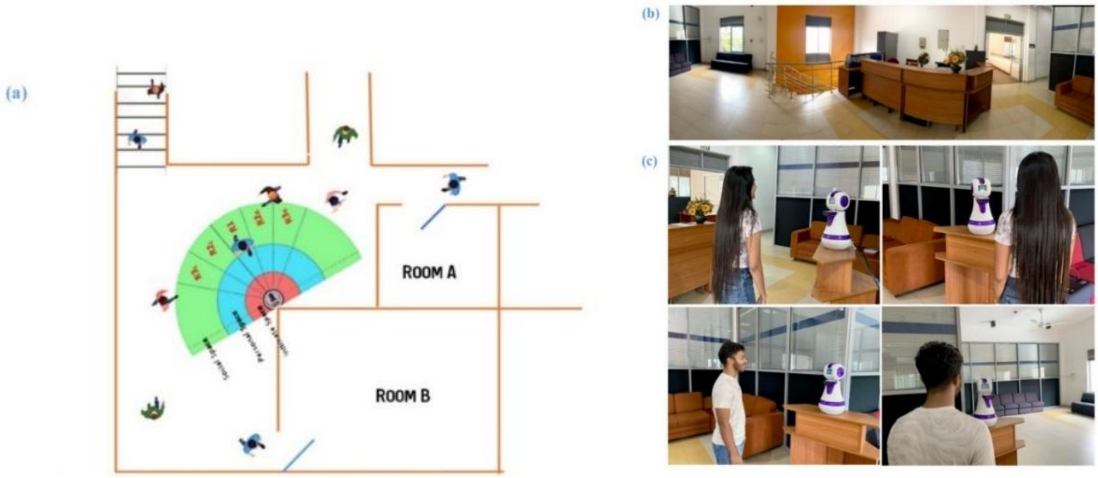


Figure 8. Experimental setup in the real world: The figure illustrates the spatial arrangement of the experiment, where the robot was positioned in an open area between Room A and Room B, serving as a common passageway. Participants approached from both the left and right sides, allowing the robot to engage with them dynamically.

a) The layout of the experimental space, highlighting the robot's position between Room A and Room B.
b) A 180-degree view of the area where the interaction takes place.
c) Participants interacting with the robot in the setup.

4.1 Interpersonal Distance Allocation and Interactive Region Assignment

Interpersonal distances followed established proxemic zones (Hall et al., 1968; Neggers et al., 2022; Weerawarna et al., 2023): intimate space (0-0.46m), personal space (0.46-1.22m), and social space (1.22-3.7m). Angular allocations aligned with human visual attention and the webcam's 120° field of view.

The Central Field of View (CFOV) covered 20° for high-acuity focus, while the Near Peripheral Fields of View (NPFOV) extended 40° each side, divided into inner (20°) and outer (20°) sections. This structure enabled smooth attention transitions:

The interactive regions correspond to specific spatial relations:

- R1 (CFOV, 20°): Directly in front, enabling immediate response.
- R2L and R2R (NPFOV, inner 20°): Near central view, requiring moderate head adjustments.
- R3L and R3R (NPFOV, outer 20°): Near field boundaries, requiring larger orientation shifts.

This segmentation allowed dynamic detection and response based on user position, creating natural human-robot interaction while minimizing abrupt movements.

*Correspondence: deepal@sjp.ac.lk

4.2 Experimental Setup and Participants

In this study, we defined five interactive regions (Table 2) and six corresponding action groups (Table 3). When a person enters the social space, the robot executes predefined action groups tailored for each region, with multiple possible actions available per region (Table 4). The robot's body rotation angle is adjusted to align with the specific region's range. For example, body rotation is typically not performed within the R1 region during the execution of action groups C and F. Conversely, when a person is located in the R3L or R3R regions, estimating their distance without rotating the robot's body becomes difficult.

When a person transitions from the social space to the personal space to interact with the robot, it signifies a successful human attention shift. Subsequently, the actions undertaken by the robot and any attempts by the person to communicate with the robot contribute to a successful human-robot interaction.

Table 3. Interactive Regions and Perfume Action Groups

Interactive regions	Action Group
R1	Action A, Action B, Action D, Action E
R2$_L$	Action A, Action B, Action C, Action D, Action E, Action F
R2$_R$	Action A, Action B, Action C, Action D, Action E, Action F
R3$_L$	Action B, Action C, Action E, Action F
R3$_R$	Action B, Action C, Action E, Action F

5. Results and Discussion

A summary of the key findings demonstrates that simple actions involving head movement and verbal cues (Action D) were most effective in the Central Region (R1). In contrast, Peripheral Regions (R2/R3) required complex, multimodal actions involving body rotation (Action F) to capture attention. Overall, Action F was the most effective universal behavior, particularly in low-engagement zones. The primary objective of this study is to identify the most effective action behaviors of the robot for capturing attention and engaging with guests across each region, while also determining the optimal human-robot interaction region. We recorded attention shifts of guests passing by based on six action groups within five regions, resulting in data collected from interactions with 228 guests. Among these interactions, successful human-robot interactions were recorded from all 228 successful human attention shifts. Table 4 presents the distribution of successful human attention shifts for the six different behaviors across the five interactive regions.

Table *5* Table 5 illustrates the distribution of successful human-robot interactions for the six different behaviors across the five interactive regions. When a person transitions from the social space to the personal space to interact with the robot, it signifies a successful human attention shift. Subsequently, the actions undertaken by the robot and any attempts by the person to communicate with the robot contribute to a successful human-robot interaction.

*Correspondence: deepal@sjp.ac.lk

Table 4. Distribution of the successful human attention shift of the six different behaviors in five different interactive regions

Interactive regions	Action Groups						
	Action A	Action B	Action C	Action D	Action E	Action F	Total
R1	9	8	X	11	13	X	**24**
$R2_L$	6	5	9	8	9	16	**53**
$R2_R$	5	7	11	7	6	17	**53**
$R3_L$	X	5	13	X	7	18	**25**
$R3_R$	X	4	11	X	6	17	**23**
Total	**20**	**29**	**44**	**26**	**41**	**68**	**228**

Table 5. Distribution of the successful human-robot interaction of the six different behaviors in five different interactive regions

Interactive regions	Action Groups						
	Action A	Action B	Action C	Action D	Action E	Action F	
R1	2	1	X	11	13	X	**24**
$R2_L$	0	2	2	8	9	16	**37**
$R2_R$	1	1	3	7	6	17	**35**
$R3_L$	X	0	3	X	7	18	**25**
$R3_R$	X	2	2	X	6	17	**23**
Total	**3**	**6**	**10**	**26**	**41**	**68**	**154**

5.1 Analysis of Successful Human Attention Shifts

To determine the optimal robot actions across different spatial regions, we conducted experiments focusing on attention shifts during human-robot interactions. These interactions were analyzed in relation to spatial factors, aligning with existing research on peripheral vision and social robotics. Observations were categorized using two nominal variables: interactive region and action type. The five spatial regions (R1, R2L, R2R, R3L, R3R) correspond to varying degrees of human proximity, while the six action categories (A–F) represent distinct robot behaviors aimed at engaging passersby. To analyze these categorical relationships, we employed correspondence analysis, a statistical method commonly used in human-robot interaction studies.

5.1.1 Contingency Table Analysis

We begin our analysis with Table 6, which shows the distribution of successful human attention shifts across different interactive regions and action types.

In total, 228 successful attention shifts were observed. Action F garnered the most attention shifts (68), followed by Action C (44) and Action E (41). Regions R2L and R2R had the highest number of attention shifts (53 each), while R3R had the least (38).

**Correspondence: deepal@sjp.ac.lk*

Table 6. Distribution of the successful human attention shift of the six different behaviors in five different interactive regions

Region	Action A	Action B	Action C	Action D	Action E	Action F	Active Margin
R1	9	8	0	11	13	0	**41**
$R2_L$	6	5	9	8	9	16	**53**
$R2_R$	5	7	11	7	6	17	**53**
$R3_L$	0	5	13	0	7	18	**43**
$R3_R$	0	4	11	0	6	17	**38**
Active Margin	**20**	**29**	**44**	**26**	**41**	**68**	**228**

5.1.2 Interpretation of Results

The correspondence analysis identified two key dimensions, explaining the relationships between spatial regions and robot action types. Dimension 1, which accounts for 93.7% of the inertia (Table 7), captures the primary interaction patterns, suggesting that a single dominant factor governs human-robot engagement across regions. Dimension 2 contributes only 5.3% of the inertia (Table 7), implying it captures subtler secondary patterns, potentially related to action variability within intermediate zones. The total chi-square value of 72.011 ($p < 0.001$) confirms a statistically significant relationship between robot actions and interaction success, reinforcing findings from prior research that suggest spatial context is a primary driver of human engagement.

Table 7. Summary of Correspondence Analysis Between Region and Action Type for Human Attention Shifts

					Proportion of Inertia		Confidence Singular Value	
Dimension	**Singular Value**	**Inertia**	**Chi Square**	**Sig.**	**Accounted for**	**Cumulative**	**Standard Deviation**	**Correlation 2**
1	0.544	0.296			0.937	0.937	0.157	0.254
2	0.129	0.017			0.053	1.000	0.044	
Total		0.316	72.011	0	1.000	1.000	1.000	

The row point analysis highlights significant regional differences in successful human-robot interactions. Region R1's high positive score (1.96, Table 8) in Dimension 1 suggests that it is strongly associated with effective engagement behaviors, aligning with previous findings on attention shifts. Regions R3L and R3R, with high negative scores (-0.97 and -0.92, respectively), indicate an inverse relationship between these peripheral zones and certain interactive actions, meaning engagement success is lower in these areas (Table 8). Intermediate regions R2L and R2R, while showing near-zero scores in Dimension 1, exhibit greater variation in Dimension 2, suggesting they are context-dependent and may be influenced by factors such as action complexity or participant movement. This variability underscores the need for adaptive robot behaviors, particularly in transition zones where human responses are less predictable.

*Correspondence: deepal@sjp.ac.lk

Table 8. Overview of Row Points (Regions) for Successful Human Attention Shifts in Correspondence Analysis

Region	Mass	Score in Dimension		Inertia	Contribution			
					of Point to Inertia of Dimension		of Dimension to Inertia of Point	
		1	2		1	2	1	2
R1	0.18	1.96	0.89	0.59	2.26	6.17	1.21	0.37
R2L	0.24	0.06	-1.01	0.03	0.01	11.08	0.07	0.35
R2R	0.23	-0.09	-1.03	0.04	0.01	10.91	0.04	0.11
R3L	0.18	-0.97	1.37	0.18	0.58	15.34	9.79	1.13
R3R	0.18	-0.92	0.42	0.16	0.50	1.41	3.91	0.07
Total	1.00			1.00				

The overview of column points for Successful Human Attention Shifts shows how each action contributes to and is represented by the dimensions. Actions A and D have high positive scores in Dimension 1 (2.34 and 1.51), aligning them strongly with R1 (Table 9). Actions C and F have negative scores in Dimension 1 (-0.90 and -0.91), associating them with R3L and R3R. Action E has a moderate positive score in Dimension 1 (0.59) and the highest score in Dimension 2 (1.36), suggesting it has a unique profile across the regions. Action B has scores close to zero in Dimension 1, indicating it's less strongly characterized by this primary dimension.

Table 9. Overview of Row Points (Regions) for Successful Human Attention Shifts in Correspondence Analysis

Region	Mass	Score in Dimension		Inertia	Contribution			
					of Point to Inertia of Dimension		of Dimension to Inertia of Point	
		1	2		1	2	1	2
Action A	0.02	2.34	1.64	0.11	0.36	2.37	5.08	0.50
Action B	0.04	0.05	-1.49	0.04	0.00	3.87	0.02	0.18
Action C	0.06	-0.90	-0.12	0.05	0.18	0.04	7.33	0.00
Action D	0.17	1.51	-1.45	0.36	1.29	15.84	3.96	0.73
Action E	0.27	0.59	1.36	0.12	0.31	22.07	2.54	0.15
Action F	0.44	-0.91	-0.19	0.32	1.22	0.72	7.13	0.01
Total	1			1				

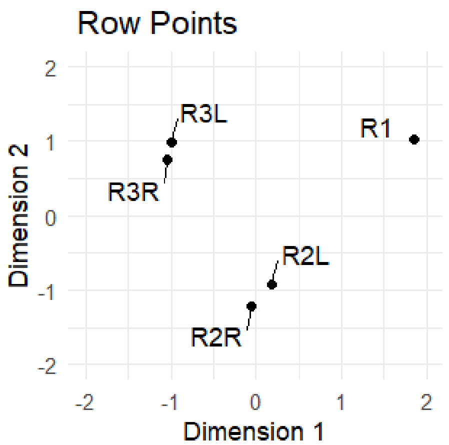
Most interactive action for each region:

- R1: Actions A (simple head movement) and D (head movement with verbal behavior)
- R2L and R2R: Action F (partial body rotation with head movement and verbal behaviors)
- R3L and R3R: Actions C (body rotation) and F

*Correspondence: deepal@sjp.ac.lk

Most interactive action overall: Action F, due to its high frequency (68 occurrences) and strong association with multiple regions.

Figure 9(a) illustrates the spatial distribution of interaction regions in a two-dimensional space, revealing distinct patterns based on proximity to the robot. The central zone (R1) is positioned in the upper right quadrant, while the intermediate zone (R2L and R2R) occupies the lower right quadrant, and the peripheral zone (R3L and R3R) is located in the upper left quadrant. Dimension 1 delineates these three main zones, while Dimension 2 provides further distinction between sub-regions within the intermediate and peripheral areas. Notably, R2L and R2R form a cluster, as do R3L and R3R, indicating similarities within these paired regions. This distribution implies distinct profiles for different spatial zones, suggesting that the effectiveness of robot actions likely varies based on the spatial relationship between the robot and human. The figure underscores the critical importance of considering spatial context in designing and implementing human-robot interactions.



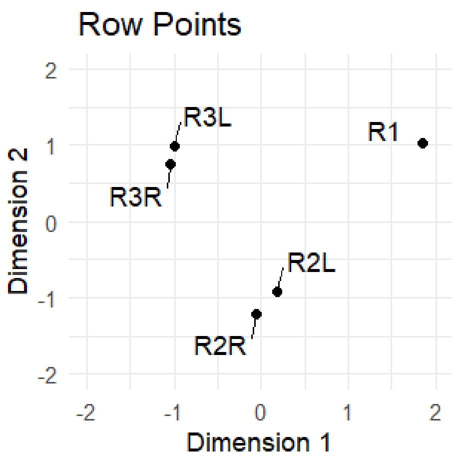


Figure 9. (a) Spatial Distribution of Interaction Regions in a Two-Dimensional Space for Successful Human Attention Shifts. (b) Positioning of Robot Actions in Terms of Effectiveness for Successful Human Attention Shifts in a Two-Dimensional Space

Figure 9(b) displays the positioning of various robot actions within the same two-dimensional space, revealing insights into their relative effects and similarities. Actions A and D cluster in the lower right quadrant, while Actions C and F show proximity on the left side of the plot. Action E occupies an isolated position in the upper right quadrant, suggesting a unique profile, whereas Action B's central location implies a more averaged effect across regions. Dimension 1 separates the actions into distinct groups (A/D, E, C/F), while Dimension 2 further distinguishes between actions within these groups. This distribution indicates that different actions have distinct effects on human attention and interaction, highlighting the potential for tailoring actions to specific interaction contexts. The diverse positioning of the actions underscores the complex and varied impact of different robot behaviors, providing valuable insights for optimizing human-robot interactions across various spatial configurations.

The comprehensive plot (Figure 10) combines regions and actions, providing insights into their relationships. The proximity of R1 to Actions A and D suggests these actions are most effective in the central region, while R3L and R3R's association with Actions C and F indicates their effectiveness in peripheral areas. Action E's central position implies moderate effectiveness across multiple regions. The visualization demonstrates that certain actions are more effective in specific spatial configurations, highlighting the potential for optimizing robot behavior based on its relative position to humans. The

*Correspondence: deepal@sjp.ac.lk

plot emphasizes the importance of adaptive behavior in human-robot interaction and provides a foundation for designing more effective interaction strategies in public spaces.

.

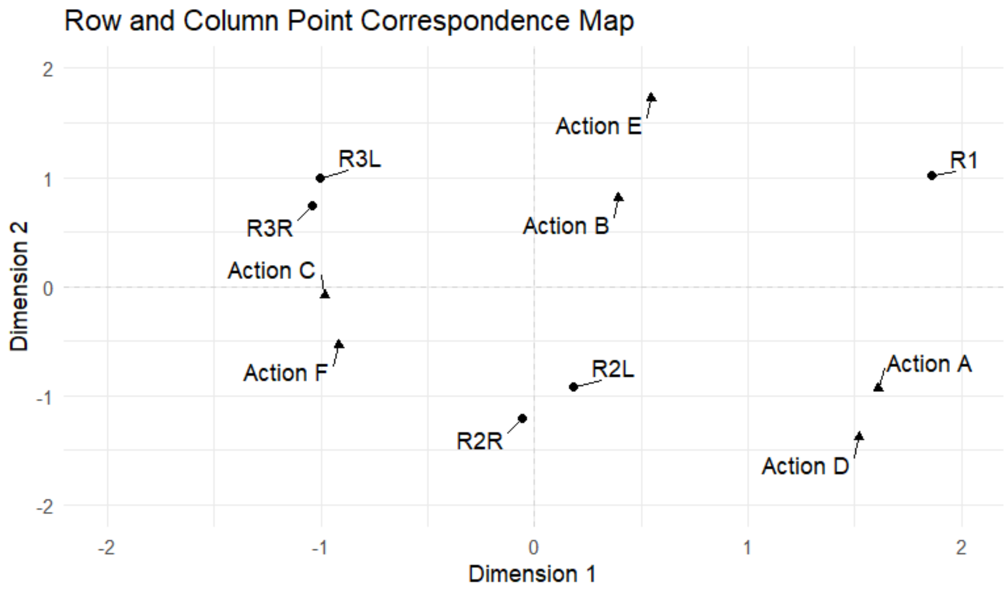


Figure10. Comprehensive Map Combining Regions and Actions for Successful Human Attention Shifts, Highlighting Interplay Between Spatial Context and Action Complexity

5.2 Analysis of Successful Human-Robot Interactions

In this section, we analyze successful human-robot interactions, focusing on cases where the interaction progressed beyond an initial attention shift to meaningful engagement. Following the same procedure as above, we explore how different robot actions contribute to engagement success across spatial regions. This analysis provides deeper insights into how spatial positioning and behavioral choices influence interaction outcomes, complementing our earlier findings on attention shifts.

5.2.1 Contingency Table Analysis

Table 10 presents the distribution of successful human-robot interactions across different regions and action types, offering insights into which behaviors are most effective in initiating engagement.

In total, 154 successful human-robot interactions were observed. Action F again garnered the most interactions (68), followed by Action E (41) and Action D (26). Region R2L had the highest number of interactions (37), while R1 and R3R had the least (27 each).

*Correspondence: deepal@sjp.ac.lk

Table 10. Distribution of the successful human-robot interaction of the six different behaviors in five different interactive regions

Interactive regions	Action Groups						
	Interactions by Initiator						Active Margin
	Person			Robot			
	Action A	Action B	Action C	Action D	Action E	Action F	
R1	2	1	X	11	13	X	**24**
$R2_L$	0	2	2	8	9	16	**37**
$R2_R$	1	1	3	7	6	17	**35**
$R3_L$	X	0	3	X	7	18	**25**
$R3_R$	X	2	2	X	6	17	**23**
Active Margin	**3**	**6**	**10**	**26**	**41**	**68**	**154**

5.2.3 Interpretation of Results

The correspondence analysis identified two key dimensions, explaining how different regions and robot actions contribute to successful human-robot interactions. Dimension 1, which accounts for 86.4% of the inertia (Table 11), primarily reflects spatial positioning effects, suggesting that the proximity of human subjects to robot plays a major role in interaction success. Dimension 2, contributing 6.4%, likely represents secondary factors such as action complexity or response timing. The total chi-square value of 53.27 ($p < 0.001$) confirms a strong statistical association between regions and robot actions, reinforcing findings from previous studies on human-robot engagement. However, compared to attention shifts, the association is slightly less pronounced, indicating that while spatiality is crucial, other factors may influence whether an interaction progresses beyond initial attention.

Table 11. Summary of Correspondence Analysis for Successful Human-Robot Interactions Across Interactive Regions and Action Groups

					Proportion of Inertia		Confidence Singular Value	
Dimension	**Singular Value**	**Inertia**	**Chi Square**	**Sig.**	**Accounted for**	**Cumulative**	**Standard Deviation**	**Correlation 2**
1	0.547	0.299	NA		0.864	0.864	0.160	-0.045
2	0.149	0.022	NA		0.064	1.000	0.045	
Total		0.346	53.27	0	1.000	1.000		

The overview of row points for Successful Human-Robot Interactions shows that R1 has the highest positive score in Dimension 1 (Table 12), similar to the attention shifts analysis. R3L and R3R have high negative scores in Dimension 1, again mirroring the attention shifts results. R2L and R2R have scores closer to zero in Dimension 1, but show more variation in Dimension 2, suggesting some nuanced differences in how these intermediate regions relate to successful interactions.

*Correspondence: deepal@sjp.ac.lk

Table 12. Overview of Column Points (Actions) for Successful Human-Robot Interactions in Correspondence Analysis

Region	Mass	Score in Dimension		Inertia	Contribution			
					of Point to Inertia of Dimension		of Dimension to Inertia of Point	
		1	2		1	2	1	2
Action A	0.02	2.34	1.64	0.11	0.36	2.37	5.08	0.50
Action B	0.04	0.05	-1.49	0.04	0.00	3.87	0.02	0.18
Action C	0.06	-0.90	-0.12	0.05	0.18	0.04	7.33	0.00
Action D	0.17	1.51	-1.45	0.36	1.29	15.84	3.96	0.73
Action E	0.27	0.59	1.36	0.12	0.31	22.07	2.54	0.15
Action F	0.44	-0.91	-0.19	0.32	1.22	0.72	7.13	0.01
Total	1			1				

The overview of column points for Successful Human-Robot Interactions reveals that Actions D and E have high positive scores in Dimension 1 (Table 13), associating them strongly with R1. Action F has a high negative score in Dimension 1, associated with R3L and R3R. Actions A, B, and C have fewer extreme scores, indicating they're less strongly characterized by the primary dimension in terms of successful interactions.

Table 113. Overview of Row Points (Regions) for Successful Human-Robot Interactions in Correspondence Analysis

Region	Mass	Score in Dimension		Inertia	Contribution			
					of Point to Inertia of Dimension		of Dimension to Inertia of Point	
		1	2		1	2	1	2
R1	0.18	1.96	0.89	0.59	2.26	6.17	1.21	0.37
R2L	0.24	0.06	-1.01	0.03	0.00	11.08	0.07	0.35
R2R	0.23	-0.09	-1.03	0.04	0.01	10.91	0.04	0.11
R3L	0.18	-0.97	1.37	0.18	0.58	15.34	9.79	1.13
R3R	0.18	-0.92	0.42	0.16	0.50	1.41	3.91	0.07
Total	1.00			1.00				

Interpreting these results, we observe:

- Region R1 is strongly associated with Actions D and E, similar to the attention shift analysis.
- Regions R2L and R2R are again closely related and associated with Action F.
- Regions R3L and R3R are strongly associated with Actions C and F.

**Correspondence: deepal@sjp.ac.lk*

- Actions A and B have less influence on successful interactions compared to attention shifts.
- Most interactive action overall: Action F, due to its high frequency (68 occurrences) and strong association with multiple regions, particularly the peripheral ones.

Figure 11(a) (row points) illustrates the distribution of interaction regions for successful human-robot interactions. R1 is isolated in the upper right quadrant, indicating a unique interaction profile in the central zone. R2L and R2R are clustered in the lower center, suggesting similar interaction patterns in these intermediate zones. R3L and R3R are positioned in the upper left quadrant, implying comparable interaction characteristics in peripheral areas. The clear separation along Dimension 1 highlights the significant impact of spatial proximity on interaction success, while Dimension 2 reveals subtle differences within similar zones. This distribution emphasizes the crucial role of spatial context in shaping human-robot interaction outcomes and suggests that robots need to adapt their interaction strategies based on their relative position to humans.

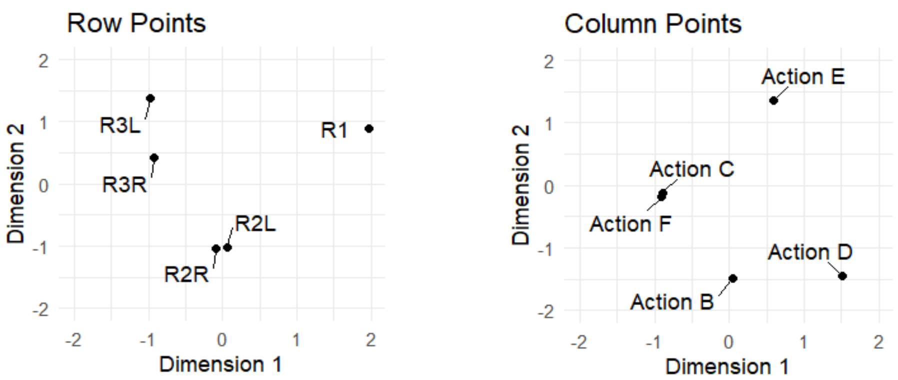


Figure 11. (a) Distribution of Interaction Regions for Successful Human-Robot Interactions in a Two-Dimensional Space. (b) Positioning of Robot Actions in Terms of Effectiveness for Successful Human-Robot Interactions in a Two-Dimensional Space.

Figure 11(b) (Column Point) displays the positioning of robot actions in terms of their effectiveness in successful human-robot interactions. Actions C and F cluster in the left quadrant, indicating similarity in their impact, possibly related to more complex body movements. Action E occupies a unique position in the upper right quadrant, suggesting a distinct effect on interactions. Actions A and D are separated along Dimension 2 but share similar positions on Dimension 1, implying some shared characteristics despite differences. Action B's central location suggests a moderate, balanced effect across different regions. This distribution underscores the diverse impacts of various robot behaviors on interaction success and highlights the potential for tailoring actions to specific spatial contexts to optimize human-robot communication.

This comprehensive map (Figure 12) combining insights from both regions and actions, offering a holistic view of successful human-robot interactions. The proximity of R1 to Actions A and D suggests these simpler actions are most effective in close-range interactions. R3L and R3R's association with Actions C and F indicates that more complex, full-body actions are more successful in peripheral zones. R2L and R2R's intermediate position, with some proximity to Action B, implies a transition zone where moderate actions are effective. Action E's central yet isolated position suggests a uniquely versatile action effective across multiple regions. This visualization demonstrates the interplay between spatial context and action complexity in determining interaction success, providing valuable insights for designing adaptive robot behaviors that can optimize interactions across various spatial configurations in public spaces.

*Correspondence: deepal@sjp.ac.lk

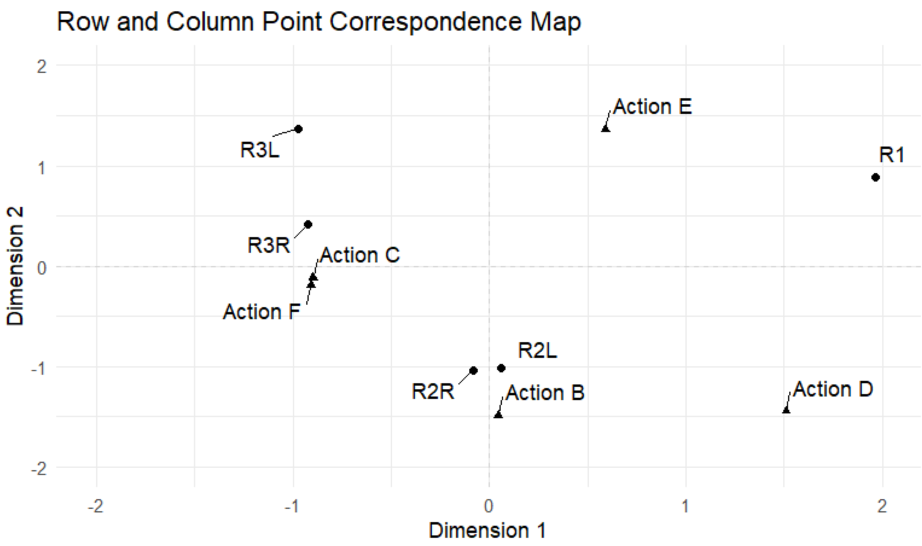


Figure 12. Comprehensive Map Combining Regions and Actions for Successful Human-Robot Interactions, Illustrating the Interplay Between Spatial Context and Action Complexity.

5.3 Comparison and Overall Conclusions

The analysis of human attention shifts and successful human-robot interactions reveals valuable insights into the effectiveness of different robot actions across various interactive regions.

Most Interactive Actions by Region:

R1 (Central Field of View): For attention shifts, Actions A (head movement) and D (head movement with verbal behavior) were most effective. However, for interactions, Actions D and E (head movement with verbal behavior) were dominant. Action D, combining head movement with verbal cues, is consistently effective in the central region. This aligns with Hall's proxemics theory, which suggests that intimate and personal spaces require less exaggerated signaling for effective communication.

R2L and R2R (Near Peripheral Fields of View): Action F (body rotation with head movement and verbal behaviors) proved the most effective for both attention shifts and interactions. Action F is highly effective in near peripheral regions, where more complex, dynamic actions capture attention.

R3L and R3R (Peripheral Fields of View): Action C (body rotation) was most successful for attention shifts, while Actions C and F proved effective for interactions. In peripheral regions, more intricate actions like Action F (combining body rotation and verbal cues) are crucial to attract attention and engage users.

Overall, Most Interactive Action: Action F (combining body rotation with head movement and verbal behavior) emerges as the most interactive action overall, showing high frequencies in both attention shifts and successful interactions. Action F proved especially effective in peripheral regions. This supports findings by Hoque et al., who noted that peripheral attention modulation requires deliberate, high-magnitude redirection signals. The broader motions of Action F successfully bridge the gap in visual attention in these outer zones.

*Correspondence: deepal@sjp.ac.lk

Key Insights: Action F is consistently the most effective in capturing and maintaining human attention across various spatial regions, especially in areas where human engagement is generally more challenging (e.g., peripheral zones).
Simpler actions like Action D, which involves head movement with verbal cues, are more effective in close proximity (R1), where interactions can be more direct and low-energy. These findings emphasize the need for robots to adapt their actions based on their spatial relationship with humans, utilizing simpler, direct actions in central zones and more complex, dynamic interactions in peripheral zones.

The consistency of these results across both attention shifts and successful interactions strengthens our understanding that spatial proximity plays a key role in robot engagement. Future designs of socially interactive robots should focus on real-time, adaptive behavioral algorithms that consider spatial context to optimize interactions. By dynamically adjusting robot behaviors based on user proximity and environmental cues, we can significantly enhance human-robot interaction efficiency and effectiveness.

6. Conclusions

Our correspondence analysis of both human attention shifts, and successful human-robot interactions highlights the critical role of spatial positioning in determining interaction effectiveness. The findings indicate that robot actions should be adapted to different spatial regions to optimize engagement. Action F (partial body rotation with head movement and verbal behaviors) proved to be the most effective overall, particularly in peripheral regions (R2L, R2R, R3L, R3R), were broader motion increases visibility and engagement. Conversely, simpler actions like head movements and verbal cues (Actions A, D, and E) were most effective in the central region (R1), where close proximity allows for more direct interaction without requiring exaggerated movements.

Multimodal actions (combining movement and speech) are particularly important in peripheral areas, where human engagement is lower, requiring more dynamic cues to capture attention. Meanwhile, simpler actions are preferable in central zones, where human-robot proximity allows for more direct, low-energy interactions. The consistency between attention shift and successful interaction results reinforces the idea that robots should dynamically adjust their engagement strategies based on spatial positioning. These insights have significant implications for the design of interactive robots in public spaces, suggesting that real-time behavior adaptation can enhance human-robot communication efficiency.

These findings establish a foundation for developing socially interactive robots that can optimize their behaviors based on spatial awareness and human responsiveness. By dynamically adjusting actions according to user proximity, these robots can enhance both engagement effectiveness and naturalness in public interactions.

Future research should explore the implementation of real-time adaptive behavioral algorithms, allowing robots to dynamically adjust interactions based on spatial and user-specific cues. Additionally, cross-cultural studies could investigate how human-robot interaction preferences vary across different societies, influencing engagement strategies.

References

Avila, L., Bailey, M., 2015. High Tech @ Home. IEEE Comput. Graph. Appl. 35, 8–9. https://doi.org/10.1109/MCG.2015.56

Breazeal, C., Dautenhahn, K., Kanda, T., 2016. Social Robotics, in: Siciliano, B., Khatib, O. (Eds.), Springer Handbook of Robotics. Springer International Publishing, Cham, pp. 1935–1972. https://doi.org/10.1007/978-3-319-32552-1_72

**Correspondence: deepal@sjp.ac.lk*

Capozzi, F., Ristic, J., 2018. How attention gates social interactions. Ann. N. Y. Acad. Sci. https://doi.org/10.1111/nyas.13854

Carraro, L., Dalmaso, M., Castelli, L., Galfano, G., Bobbio, A., Mantovani, G., 2017. The appeal of the devil's eye: Social evaluation affects social attention. Cogn. Process. 18, 97–103. https://doi.org/10.1007/s10339-016-0785-2

Das, D., Kobayashi, Y., Kuno, Y., 2013. Attracting attention and establishing a communication channel based on the level of visual focus of attention, in: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. Presented at the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013), IEEE, Tokyo, pp. 2194–2201. https://doi.org/10.1109/IROS.2013.6696663

Datta, C., Vijay, R., 2010. Neel: An intelligent shopping guide using web data for rich interactions, in: 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI). Presented at the 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 87–88. https://doi.org/10.1109/HRI.2010.5453256

DiSalvo, C.F., Gemperle, F., Forlizzi, J., Kiesler, S., 2002. All robots are not created equal: the design and perception of humanoid robot heads, in: Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS '02. Association for Computing Machinery, New York, NY, USA, pp. 321–326. https://doi.org/10.1145/778712.778756

Iwasaki, M., Ogawa, K., Yamasaki, A., Yamazaki, K., Miyazaki, Y., Kawamura, T., Nakanishi, H., Enabling Shared Attention with Customers Strengthens a Sales Robot's Social Presence, International Conference on Human Agent Interaction 2022 DOI:10.1145/3527188.3561918

Frijns, H.A., Schürer, O., Koeszegi, S.T., 2023. Communication Models in Human–Robot Interaction: An Asymmetric MODel of ALterity in Human–Robot Interaction (AMODAL-HRI). Int. J. Soc. Robot. 15, 473–500. https://doi.org/10.1007/s12369-021-00785-7

Glas, D.F., Kamei, K., Kanda, T., Miyashita, T., Hagita, N., 2015. Human-Robot Interaction in Public and Smart Spaces, in: Mohammed, S., Moreno, J.C., Kong, K., Amirat, Y. (Eds.), Intelligent Assistive Robots: Recent Advances in Assistive Robotics for Everyday Activities. Springer International Publishing, Cham, pp. 235–273. https://doi.org/10.1007/978-3-319-12922-8_9

Gobel, M.S., Giesbrecht, B., 2020. Social information rapidly prioritizes overt but not covert attention in a joint spatial cueing task. Acta Psychol. (Amst.) 211, 103188. https://doi.org/10.1016/j.actpsy.2020.103188

Hall, E.T., Birdwhistell, R.L., Bock, B., Bohannan, P., Diebold, A.R., Durbin, M., Edmonson, M.S., Fischer, J.L., Hymes, D., Kimball, S.T., Barre, W.L., Lynch, F., J., S., McClellan, J.E., Marshall, D.S., Milner, G.B., Sarles, H.B., Trager, G.L., Vayda, A.P., 1968. Proxemics [and Comments and Replies]. Curr. Anthropol. 9, 83–108.

Hayward, D.A., Voorhies, W., Morris, J.L., Capozzi, F., Ristic, J., 2017. Staring reality in the face: A comparison of social attention across laboratory and real world measures suggests little common ground. Can. J. Exp. Psychol. Rev. Can. Psychol. Expérimentale 71, 212–225. https://doi.org/10.1037/cep0000117

Hegel, F., Muhl, C., Wrede, B., Hielscher-Fastabend, M., Sagerer, G., 2009. Understanding Social Robots, in: 2009 Second International Conferences on Advances in Computer-Human Interactions. Presented at the 2009 Second International Conferences on Advances in Computer-Human Interactions, pp. 169–174. https://doi.org/10.1109/ACHI.2009.51

Hoque, M.M., Das, D., Onuki, T., Kobayashi, Y., Kuno, Y., 2012a. Robotic System Controlling Target Human's Attention, in: Huang, D.-S., Ma, J., Jo, K.-H., Gromiha, M.M. (Eds.), Intelligent Computing Theories and Applications. Springer, Berlin, Heidelberg, pp. 534–544. https://doi.org/10.1007/978-3-642-31576-3_68

*Correspondence: deepal@sjp.ac.lk

Hoque, M.M., Das, D., Onuki, T., Kobayashi, Y., Kuno, Y., 2012b. An integrated approach of attention control of target human by nonverbal behaviors of robots in different viewing situations. 2012 IEEERSJ Int. Conf. Intell. Robots Syst. 1399–1406. https://doi.org/10.1109/IROS.2012.6385480

Hoque, M.M., Kobayashi, Y., Kuno, Y., 2014. A Proactive Approach of Robotic Framework for Making Eye Contact with Humans. Adv. Hum.-Comput. Interact. 2014, 694046. https://doi.org/10.1155/2014/694046

Kwak, S.S., 2014. The Impact of the Robot Appearance Types on Social Interaction with a Robot and Service Evaluation of a Robot. Arch. Des. Res. https://doi.org/10.15187/adr.2014.05.110.2.81

Masse, B., Ba, S., Horaud, R., 2018. Tracking Gaze and Visual Focus of Attention of People Involved in Social Interaction. IEEE Trans. Pattern Anal. Mach. Intell. 40, 2711–2724. https://doi.org/10.1109/TPAMI.2017.2782819

Neggers, M.M.E., Cuijpers, R.H., Ruijten, P.A.M., IJsselsteijn, W.A., 2022. Determining Shape and Size of Personal Space of a Human when Passed by a Robot. Int. J. Soc. Robot. 14, 561–572. https://doi.org/10.1007/s12369-021-00805-6

Wang, S., Lilienfeld, S.O., Rochat, P., 2015. The uncanny valley: Existence and explanations. Rev. Gen. Psychol. 19, 393–407. https://doi.org/10.1037/gpr0000056

Weerawarna, N.T., Manawadu, U., Silva, P.R.S.D., 2023. Sociable Robot 'Lometh': Exploring Interactive Regions of a Product-Promoting Robot in a Supermarket. J. ICT Res. Appl. 17, 58–81. https://doi.org/10.5614/itbj.ict.res.appl.2023.17.1.5

Yoshiike, Y., Yamaji, Y., Miyake, T., De Silva, P.R.S., Okada, M., 2010. Sociable trash box, in: Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, HRI '10. IEEE Press, Osaka, Japan, pp. 337–338.

**Correspondence: deepal@sjp.ac.lk*