Tech Science Press

Behavior of Delivery Robot in Human-Robot Collaborative Spaces During Navigation

Kiran Jot Singh[1], Divneet Singh Kapoor[1,*], Mohamed Abouhawwash[2,3], Jehad F. Al-Amri[4], Shubham Mahajan[5] and Amit Kant Pandit[5]

[1]Embedded Systems & Robotics Research Group, Chandigarh University, Mohali, 140413, Punjab, India
[2]Department of Mathematics, Faculty of Science, Mansoura University, Mansoura, 35516, Egypt
[3]Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824, USA
[4]Department of Information Technology, College of Computers and Information Technology, Taif University, Taif, 21944, Saudi Arabia
[5]School of Electronics and Communication, Shri Mata Vaishno Devi University, Katra, 182320, India
*Corresponding Author: Divneet Singh Kapoor. Email: divneet.singh.kapoor@gmail.com

Abstract: Navigation is an essential skill for robots. It becomes a cumbersome task for the robot in a human-populated environment, and Industry 5.0 is an emerging trend that focuses on the interaction between humans and robots. Robot behavior in a social setting is the key to human acceptance while ensuring human comfort and safety. With the advancement in robotics technology, the true use cases of robots in the tourism and hospitality industry are expanding in number. There are very few experimental studies focusing on how people perceive the navigation behavior of a delivery robot. A robotic platform named "PI" has been designed, which incorporates proximity and vision sensors. The robot utilizes a real-time object recognition algorithm based on the You Only Look Once (YOLO) algorithm to detect objects and humans during navigation. This study is aimed towards evaluating human experience, for which we conducted a study among 36 participants to explore the perceived social presence, role, and perception of a delivery robot exhibiting different behavior conditions while navigating in a hotel corridor. The participants' responses were collected and compared for different behavior conditions demonstrated by the robot and results show that humans prefer an assistant role of a robot enabled with audio and visual aids exhibiting social behavior. Further, this study can be useful for developers to gain insight into the expected behavior of a delivery robot.

Keywords: Human-robot interaction; robot navigation; robot behavior; collaborative spaces; industrial IoT; industry 5.0

1 Introduction

One of the emerging themes of Industry 5.0 is human and robot collaboration and it is significantly going to affect the way we conduct our businesses. This wave of Industry 5.0 is going to define the rules and systems for human-robot interaction [1]. Robots will be moving beyond structured environments like manufacturing plants and will be entering into the everyday lives of humans e.g., hospitals, offices, homes, hotels, etc. in the form of service robots. Service robots are normally considered for a human-populated environment where they share space with humans. Most service robots are semi-self-ruling or completely self-sufficient robots with partial or full mobility. Almost 82,000 robots in the logistic domain were sold in the year 2018 which is 2.15 times more than the year 2017. Similarly, for the public relations domain, almost 33000 robots were sold in the year 2018 which is 3.3 times more than the year 2017 [2,3]. This is quite evident that the sale of robots has increased over the years and one of the factors for this surge in sale for logistic and public relation robots is the growing adoption by hotel and tourism industry.

People from different countries of the world were asked about the appropriateness of the use of robots in different domains of the tourism and hotel industry, which led to the calculation of perceived appropriateness [4–6]. From the point of view of a customer, robots are service operators that convey services to them. They have various perspectives on robots and eagerness to-pay for robot-delivered services. Besides, customers have various inclinations for services, exercises, and errands that are appropriate for robotization, modeled by their perspectives towards robots. Their attitude can be impacted by parameters like gender, occupation, awareness, etc. These inclinations impact customer goals to utilize the services of robots in the tourism and hospitality industry. People from different countries of the world were asked about the appropriateness of the use of robots in different domains, which lead to the calculation of perceived appropriateness. Fig. 1 the results of the different studies of people from different parts of the world about their perceptions towards robots' services in tourism, hospitality, and travel domains.
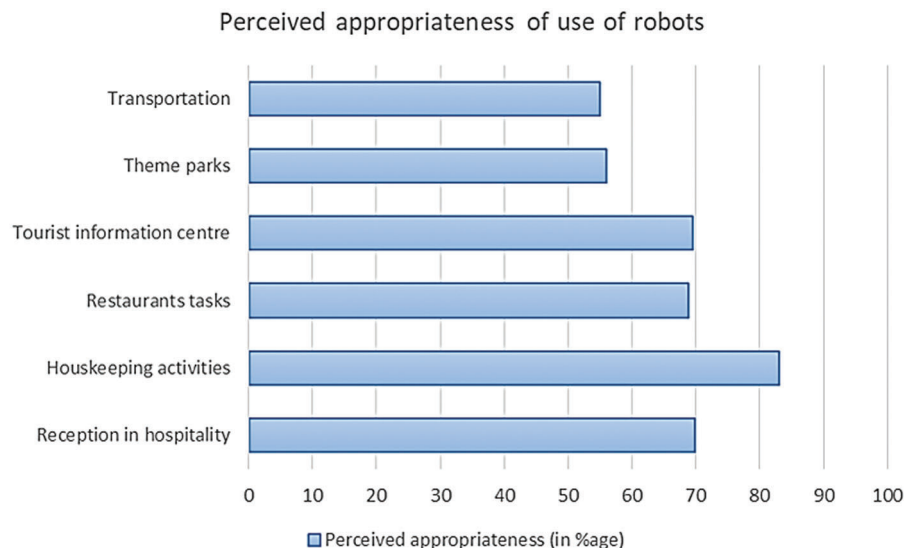

Figure 1: Perceptions towards robot's services in different domains of tourism and hospitality industry

Reception in hospitality includes different services like check-in, check-out, or guiding the room. The above results show that robots at reception are almost 70% appropriate. Housekeeping activates are composed of various activities like delivery, laundry, taking orders from customers, cleaning the rooms,

ironing, etc. People find robots to be 83% appropriate for these jobs. Restaurant tasks can be cooking food, serving food and drinks, cleaning tables, taking orders, providing information, and delivering food. Robots are considered 68% appropriate for offering these services at restaurants. Tourist Information Centers may include services like robot guiding for destinations or providing the desired information about offers. People consider robots to be almost 69% appropriate for these jobs. Theme parks can offer services like rides automation or robots controlling the rides. People find robots to be 57% appropriate for these jobs. Last is transportation which includes services like self-driving cars, buses, trains, or planes. People consider robots to be almost 56% appropriate for these jobs.

Delivering items to guest rooms came out as the most appropriate application of robots in hotels among other applications in housekeeping activities. Delivering items by a robot may involve different activities like delivering towels, linen, laundry, food, drinks, etc. One of the critical aspects of item delivery by robots in a human-populated environment will be the mobility of robots. In human-populated environments mobility of robots involves sharing the same physical space, engaging in social interactions with humans, and following the social norms. Moreover, acceptance of the robots will also depend upon factors like trust, safety, and likability of robots among humans [7–10]. So, the most vital aspect in developing robots for human-populated environments is the integration of strategies that allows them to navigate in a socially acceptable way.

Traditional robot navigation strategies focus on the metrics of execution and path planning with quantitative parameters like path length, time, smoothness, etc. Recent navigation strategies [11–16] have started focusing on socially aware robot navigation (SARN) which includes robot behavior and quality of interaction with humans. Furthermore, implementation and assessment of social component in navigation strategies necessitate the user studies, to identify the most appropriate behavior of a robot to facilitate tasks like item delivery in hotels. Behavior cues are directly linked to social cues, which refer to body movements, gestures, facial expressions, and verbal communication. The combination of all these social cues from social signals is further interpreted as emotions, personalities, or intentions [17]. A robot can make the use of behavioral cues to exhibit its mental state, which will help the robot elevate its social presence.

1.1 Motivation

As more and more robots are entering the consumer market which is meant to be part of our environment and share space with humans, it becomes increasingly important to study and identify acceptable robot social behavior. This can be achieved by introducing and studying social concepts like appropriate social distance, acknowledgment, etc. within robot navigation planning. The approach proposed in this article motivated us to find answers to the following research questions:

- What kind of role do people perceive for different behavior conditions of a delivery robot in the hotel corridor?
- What kind of behavior is preferred by humans for a delivery robot in a hotel setting?
- Do the robot's display and audio capabilities affect the perceived social presence of a delivery robot?

1.2 Contribution

It is quite evident that robots equipped with different sensors and cameras are going to be part of our society soon. In turn, it becomes increasingly important to address higher-level tasks such as navigation in human-robot collaborative spaces. In order to derive inferences, a set of three experiments have been conducted by a robot "PI" acting as a delivery robot in a hotel corridor scenario with 36 participants to cater to the following objectives:

- To find the perception of a delivery robot during navigation tasks in a hotel setting by collecting responses from human participant for different robot behaviors.
- To study the effect of audio and display capabilities of a robot on the perceived social presence of the robot.
- To study the effect of different behavior conditions on the perceived role of a robot.

2 Related Work

The preliminary attempts were made in the form of developing measurement tools proposed by Bartneck et al. [18], where they tried to measure 5 key aspects of the human-robot-interaction (HRI) i.e., animacy, likeability, anthropomorphism, perceived safety and perceived intelligence. This series was called the Godspeed questionnaire series which was intended to help the robot creators in their development journey. Mavrogiannis et al. [19] studied the effect of robot navigation strategies on human behavior, where they performed a study in a controlled lab environment for three different robot navigation strategies. They found that human acceleration reduces around an autonomous robot, as compared to a tele-operated one.

In a recent study conducted by Vassallo et al. [20], they evaluated how humans behave when they cross a robot. It was observed that human-human avoidance is quite similar to human-robot avoidance if robot follows the human interaction rules. On the same lines, Burns et al. [21] proposed a framework to promote social engagement, where they conducted an experimental study. The participants in the study exhibited a higher level of engagement and the affirmative emotional state towards the robot which displayed imitation. Rossi et al. [22] investigated the effect on trust of humans, due to the behavior of the robot. They found that people like the robots that exhibit social behaviors such as navigating at safe distance and verbally communicating with humans. Furthermore, Warta et al. [23] conducted a study on hallway navigation tasks to understand the effects of social cues on HRI. The study results indicated that human participants' perception of the social presence of a robot is highly influenced by robot proxemics behavior and robot display screen characteristics.

In the domain of mobile robotics Chatterjee et al. [24], utilized YOLO for detecting objects for a robot in multiple indoor experiment scenarios. One of the problems in robotics is limited computation capability, which was addressed by Luo et al. [25], were developed a Raspberry PI-based robot to perform real vision tasks and object detection. On similar lines, Reis et al. [26], used a single-board computer on a robot with proximity sensors powered with a YOLO algorithm to detect different obstacles and humans in an indoor environment.

As we can see in the above-listed work that researchers have working in the domain of social robotics and they are also considering vision algorithms to empower the same. It should be also noted that there human-robot interactions in an industrial environment and human-robot interaction in social spaces are two different things [27–29]. Very few studies have been conducted in social spaces to study human robot interaction. As the hotel industry is going to be early adopters of this technology, hence, it is very important to evaluate the effect of robot behavior on humans based on the natural acceptance and trust of humans. In this article, human responses have been recorded for different robot behaviors to understand their acceptance and trust towards the robot.

3 Proposed Work

This section describes the proposed approach used for identifying participants' perceptions towards the robot's role, behavior, and social presence. The complete overview of the system and approach is represented in Fig. 2. The approach has been divided into six blocks. The robot is powered by YOLO to detect humans

and obstacles during navigation. The role of the planning block is to execute different behavior conditions as described in the section below. The behavior block has three defined sets of behaviors based upon which control block is executing the instructions. The control block is receiving the data from sensors (proximity and camera) and instructing the actuators (motors, display, and speakers) to realize a specific behavior condition.
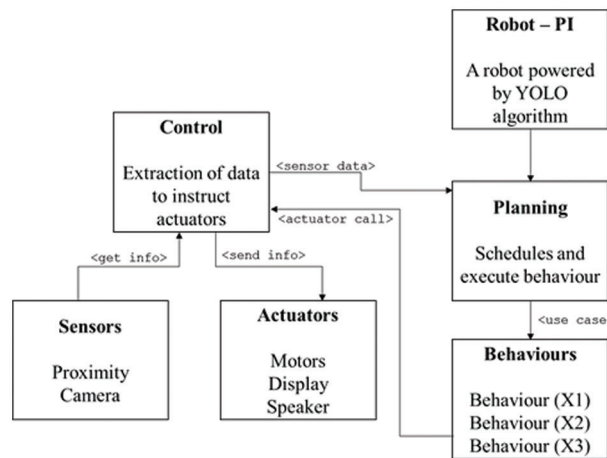

Figure 2: Complete system overview and approach

Further, the robot utilizes the above information in conjunction with proxemics constraints while navigating through the corridor space. The personal space related to individuals can be defined as the minimum space which individuals maintain to avoid discomfort. Classification of space around an individual is shown in Fig. 3.
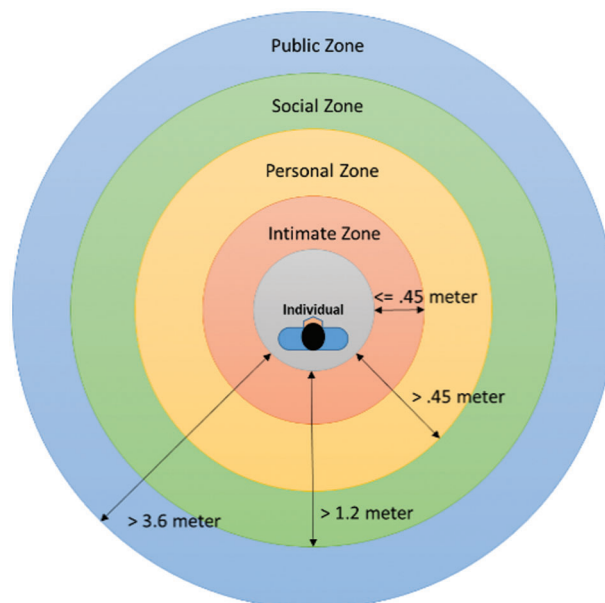

Figure 3: Classification of personal space [30]

3.1 PI Robot

We used the "PI" robotic platform as shown in Fig. 4, for conducting this study. PI is a 150 cm tall robot that consists of a screen mounted on the top of 4 wheeled platforms. PI includes a 7" LCD IPS screen with inbuilt speakers which enables it to make use of graphical and audio aids.

Figure 4: PI robotic platform used for experimentation

PI is equipped with a single-board computer (Quad-core Cortex-A72 processor, 4 GB RAM, and 2.4 GHz and 5.0 GHz IEEE 802.11ac wireless connectivity) for performing computations. PI has proximity sensor (TF-Luna, a single-point ranging LiDAR with distance range of 8 m) and a camera (Pixy 2.1 with 1296×976 resolution with integrated image flow processor) for detecting obstacles and conducting navigation. The camera is positioned on the top of LCD screen and python scripts are utilized to enable various functions like obstacle detection, vision processing, navigation etc. PI is used as a delivery robot to study the effect of different behavior conditions during navigation on humans.

3.2 YOLO for PI

PI is assumed to be navigating in a hotel corridor space shared with humans. It is assumed that PI's odometer is perfect and humans are going to avoid the robot. We model the PI as a three-dimensional system with the following dynamics:

$$x = vcos\varnothing, \ y = vsin\varnothing$$

where, $z_t = (x_i, \ y_i, \ \varnothing_i)$ is the robot state, $p_i = (x_i, \ y_i)$ is the position, $\varnothing_i$ is orientation/heading of robot, v_i is the speed of the robot. The PI observes the environment through a front facing RGB camera mounted at a fixed pitch and height. The perceptron model of PI leverages You Only Look Once (YOLO) [26,31] algorithm to predict humans and objects in the scene and where they are present. YOLO makes the use of convolutional and residual layers to detect the object and predict a bounding box around the detected object. It divides the input image from the camera into $K \times K$ grid, and one bounding box anchor is assigned for each Ground Truth (GT) object in the image. For each bounding box, the YOLO network

predicts (t_x, t_y, t_w, t_h), which is used to find the center-point (b_x, b_y), width b_w and height b_h of the predicted bounding box, given as

$$b_x = \sigma(t_x) + c_x$$

$$b_y = \sigma(t_y) + c_y$$

$$b_w = p_w exp(t_w)$$

$$b_h = p_h exp(t_h)$$

The overlap of the YOLO's predicted bounded box (PD) and GT box, known as Intersection-over-Union (IoU) is given as

$$IoU_{PD}^{GT} = \frac{area(GT \cap PD)}{area(GT \cup PD)}$$

where, $area(.)$ indicates the region. Each grid cell confidence score is associated with the bounding boxes. The confidence score depicts how much the model is confident about the accuracy of the object detected. Formally, the confidence is defined as

$$C = Pr(Obj) \times IoU_{PD}^{GT}$$

where, the confidence score is zero if no object exists in bounding boxes and otherwise confidence score should be equal to IoU_{PD}^{GT}.

The YOLO network mainly utilizes 1×1 and 3×3 convolutional kernels. The input image is processed through YOLO structure after pre-processing and image resolution adjustment. The first cluster consists of 2 convolutional layers, the output of which is given to the second cluster of 3 convolutional and 1 residual layer. The third cluster is made up of 5 convolutional and 2 residual layers, with fourth cluster comprising of 17 convolutional and 8 residual layers. The fifth cluster also consists of the same structure as of fourth cluster, and the sixth cluster is composed of 8 convolutional and 4 residual layers. Finally, the seventh cluster is composed of 3 different prediction networks that predict the different bounding box offsets.

The training of YOLO network is performed using error back propagation algorithm, in which the error between PD and GT is calculated with the aid of a loss function. The errors are utilized to adjust the network parameters to train the model towards the direction of minimization of the error and/or loss function. The loss function of YOLO algorithm is given as

$$E_{loss} = E_{localization} + E_{class} + E_{confidence}$$

$$E_{localization} = \sum_{i=1}^{N_{box}}[(b_{x,i}^{GT} - b_{x,i}^{PD})^2 + (b_{y,i}^{GT} - b_{y,i}^{PD})^2 + (b_{w,i}^{GT} - b_{w,i}^{PD})^2 + (b_{h,i}^{GT} - b_{h,i}^{PD})^2]$$

$$E_{class} = \sum_{i=1}^{N_{box}}[-\{class_i^{GT} \times log(class_i^{PD})\} - \{(1 - class_i^{GT}) \times log(1 - class_i^{PD})\}]$$

$$E_{confidence} = \sum_{i=1}^{N_{box}}[-\{C_i^{GT} \times log(C_i^{PD})\} - \{(1 - C_i^{GT}) \times log(1 - C_i^{PD})\}]$$

where, N_{box} is the maximum number of bounding boxes predicted by the YOLO algorithm.

The system utilizes the improved algorithm of the YOLO v3 model that initialize the network with key parameters, such as the convolutional layer's kernel function being set to 3×3, the stride size being 1 and the batch size being 64 for each batch of training samples. The weight decay of regularization coefficient kept at 0.0005, the adaptive learning rate being changed dynamically which is initiated from 0.01 and the maximum number of iterations kept to 2000.

During the training process, the accuracy of the three algorithm models i.e., YOLO v3 [26], Faster Region-based Convolutional Neural Network (R-CNN) [32] and Single Shot Detector (SSD) [33] in recognizing the training samples (6525) is counted when the number of iterations is 500, 1000, 1500, and 2000, respectively, to verify the detection. The statistics in Fig. 5 show that as iteration times grow, so does the average recognition accuracy of the three algorithms, with the YOLO v3 algorithm having the greatest recognition accuracy, at 94.98 percent. Further, Fig. 6 depicts the recognition results of the Person, Door, Plant, Chair and TV screen by square bounded boxes in the image. Fig. 6 clearly shows that the technique used in this study can properly find the desired elements in the image while navigation.
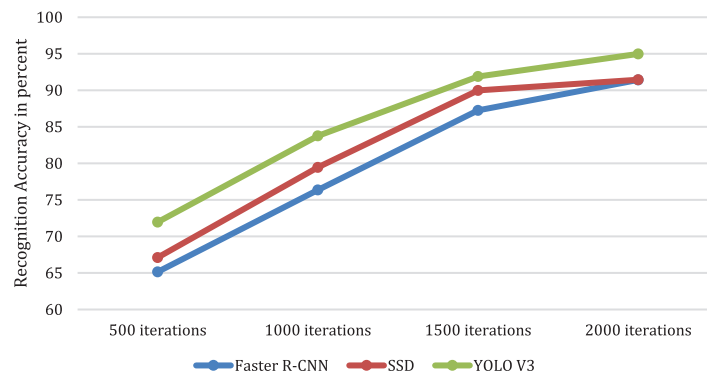

Figure 5: Recognition accuracy comparison of different object detection algorithms

Figure 6: Object detection results from real-time execution of YOLO v3 algorithm

4 Experiment

The experiment was conducted with participants ($N = 36$), ages ranging from 18 to 38. In order to test the research questions, a within-subject and repeated measure experiment design were done where the

experiment was executed for three different behavior conditions X1, X2, and X3 as given in Tab. 1. Furthermore, PI also used display and audio system as a part of behavior condition (X3) as depicted in Tab. 2. Further, to analyze the interactions between the robot and human participants, participants were asked to fill the questionnaire. This experiment is designed for participants in a hotel corridor scenario where they share the space with a delivery robot. A couple of participants were asked to position themselves in the middle of the corridor. Three participants were asked to position themselves at a turn of the corridor, in such a way that they are blocking the robot's path. One of the participants was asked to start moving in the same direction robot was moving. All participants started with behavior condition X1, then X2 and finally subjected to behavior condition X3 as illustrated in Fig. 7.

Table 1: Behaviour conditions

Condition	Movement	Obstacle	Display and audio
X1	• Moves at a constant speed	• Stop and waits for participants to clear the path if finds path blockage	• Display: OFF • Audio: OFF
X2	• Moves at the same speed as participants • Moves along with participants if moving in the same direction	• Performs obstacle avoidance if it is way-out • Stop and waits for participants to clear if the path is completely blocked	• Display: OFF • Audio: OFF
X3	• Moves at a slower speed as compared to participants • Follow the participants if moving in the same direction	• Performs obstacle avoidance if it is way-out • Ask participants to pass if the path is completely blocked	• Display: Show graphics • Audio: Turned on during asking and thanking process

Table 2: Display and audio status for different conditions

Condition	State	Display graphics	Audio message
X1	Not applicable	Not applicable	Not applicable
X2	Not applicable	Not applicable	Not applicable
X3	Asking participants to pass if the path is completely blocked	🙂	Excuse me, can I pass?
X3	Thanking participants when they clear the path	😌	Thank you!

(a)   (b)   (c)
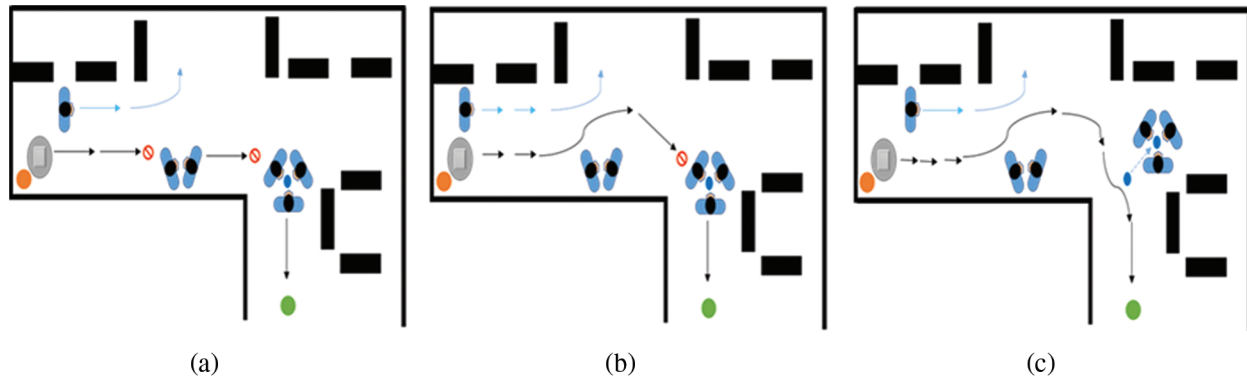
Figure 7: (a) Behavior condition X1 (B) Behavior condition X2 (C) Behavior condition X3

During X1 as shown in Fig. 7a, PI moved at a constant speed by starting to move behind the first participant navigating in the same direction. Further, it encounters two participants blocking its path as given in Tab. 1 for condition X1. So, PI simply stops and waits for participants to clear the path. Next PI again finds three participants completely blocking its way. PI waits again for the participants to clear the path so that PI can deliver the package to the desired location represented by a green circle at the end of the path. In X2 as depicted in Fig. 7b, PI starts moving along with the participant with the same speed. PI encounters the first couple of participants on the way and passes them by finding an alternate way. Further PI reaches the point where the path is completely blocked by three participants. PI waits for participants to clear the path, so that package can be delivered to the desired location. Finally, in X3 as given in Fig. 7c, PI starts following the first participant at a relatively lower speed. PI comes across the first couple of participants and takes a small turn to pass then. PI utilizes its vision procession capabilities to detect obstacles and further makes use of navigation scripts to make the decisions. As, it encounters the second group of people completely blocking the way. Now, PI as per Tab. 2 uses its audio capabilities and asks participants to give it some space to pass. PI also displays different emotions, as shown in Tab. 2 while asking and passing the participants. PI also thank participants by again using its audio capability for giving them space to pass.

At the end of every behavior condition, participants filled a questionnaire about research questions, which was further used to compute and analyze the results. Further, Fig. 8 shows a few snippets of a real scenario where the study was conducted with participants.

Figure 8: Snippets from the study conducted

5 Results

As a part of our study, participants were asked questions to find the perception of the robot at the end experiment for every behavior condition. We used the Godspeed questionnaire [8] to evaluate the perception of robots among participants for different behavior conditions. The questions targeted three different aspects of robot behavior perception. First is the likeability of a robot on a five-point scale ranging from dislike to like and unfriendly to friendly, second is the robot's perceived intelligence on a scale ranging from incompetent to competent and irresponsible to responsible, and third being participant's perceived safety on a scale ranging from of anxious to relaxed and agitated to calm for three behavior conditions of the robot. We applied the MANOVA test [34–36] to find out the effect of the robot's different behavior conditions on the participant, where p-value came out to be significant for Wilk's test (p=.036), Hoteling-Lawley's test (p=.008), and Roy's test (p=.001) with 95% confidence as shown in Tab. 3.

Table 3: P-Value for different tests

Test	P-value
Wilk's test	.036
Hotelling-lawley's test	.008
Roy's test	.001

All the tests suggested that behavior conditions have a significant effect on likeability, perceived intelligence, and perceived safety. Further, we calculated the mean score for all the dependent parameters i.e., likeability, intelligence, and safety for X1, X2, and X3 respectively, as shown in Fig. 9. The mean score for X3 for all dependent variables is observed to be higher as compared to X1 and X2. This suggests that participants liked the third behavior condition of a delivery robot during navigation in a hotel setting.
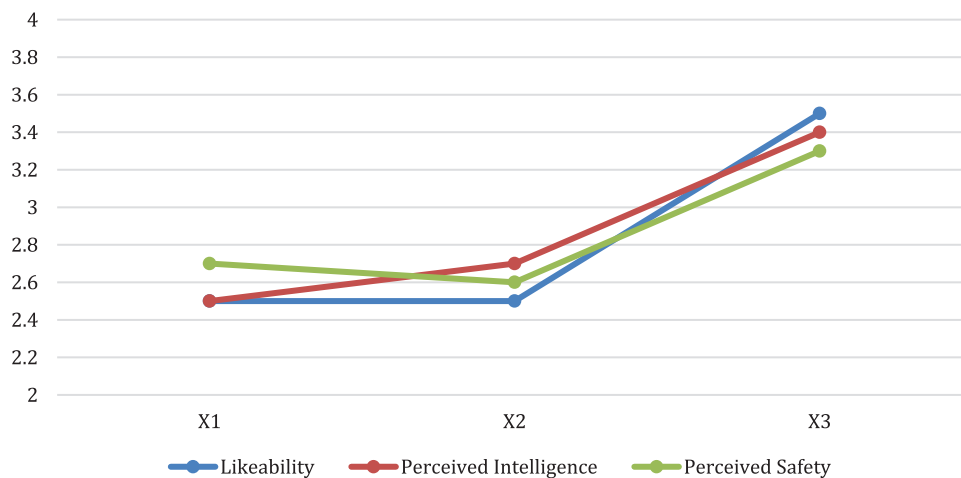

Figure 9: Behaviour condition comparison based on user reviews

As it is quite evident from the results that participants liked a robot that more, which moved with a slower speed than humans as in X3. The further robot asked participants for a path if it was completely blocked. The robot also made use of audio and visual aids which enabled the robot to exhibit its internal

state in a better way. On the whole, the robot exhibited better social behavior in X3 which made participants like it over other behavior conditions i.e., X1 and X2.

Also, every participant was asked to choose a role from 5 different role choices i.e., machine, assistant, companion, butler, and helper for the robot after the experiment for every behavior condition. Fig. 10 depicts the perceived robot role by all participants for the three behavior conditions of the robot. The robot was perceived more as a machine in X1 and as an assistant in X3. The robot has been perceived as a companion in X2, as it was moving alongside the participant at the same speed. While in X3, it followed the participant by relatively moving at lower speed and also asked for permission from the participants to pass; hence, it has been perceived more as an assistant. In X1, the robot did not showcase any social behavior, so it was perceived as a machine.
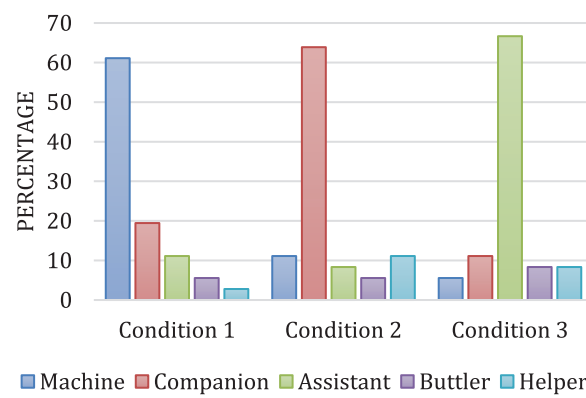

Figure 10: Perceived robot role in different behaviour conditions

As a part of our questionnaire, we also asked the participants about their preference of specific social behavior entities, they look in a robot during navigation task. They were asked to respond on a 5-point Likert scale, where 1 represented "not at all" and 5 represented "very much". The entities on which their preference was asked were: maintain an appropriate distance from humans, approach speed of the robot, and body orientation. The results for the same are depicted by Fig. 11, where it can be inferred that maintaining appropriate distance from humans is the most preferred entity among approach speed of the robot and its body orientation.
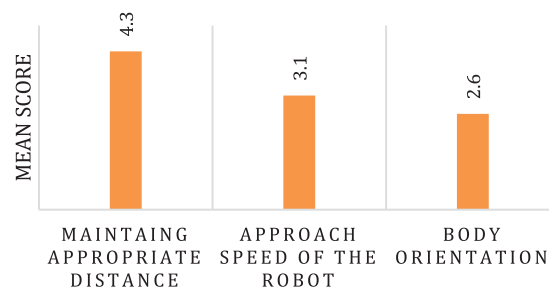

Figure 11: Participant responses on their preference of a specific social behavior entity

Further, at the end of the complete experiment, we also wanted to know about the effect of robot display and audio on perceived social presence. So, we also asked the participants the following questions as a part of the questionnaire at the end of every behavior activity.

- Does inclusion of the display and audio capabilities make a robot more sociable?
- Did you like the inclusion of display and audio features in the robot?

We asked the participants to respond on a 5-point Likert scale, where 1 represented "not at all" and 5 represented "very much". Participants supported the thought that the display capabilities have a significant effect on the robot's social presence and participants also liked the idea of inclusion of the display and audio features as depicted in Tab. 4. It can be inferred from the participant's responses that the inclusion of features like display and audio, raises the degree of likeliness towards the robot, which leads to enhanced social presence.

Table 4: Participant responses on effect of robot display and audio on perceived social presence

Question	Max	Min	Mean	Standard deviation
Does the inclusion of the display and audio capabilities make a robot more sociable?	4	3	3.6	0.48
Did you like the inclusion of display and audio features in the robot?	5	3	4.2	0.63

6 Conclusion

Navigation is one of the crucial aspects for the implementation of effective human-robot interaction. As the use of mobile robots is increasing in social settings, it becomes very important to understand the effect of different behavior conditions of a robot on humans. A robotic platform "PI" was designed utilizing the YOLO algorithm to conduct a study on the perception of a delivery robot during navigation tasks. This paper assesses and compares participant responses to explore the perceived social presence, role, and perception of a delivery robot in a hotel corridor setting performing a navigation task. The robot followed various human navigation traits i.e., moving at different speeds, passing humans at a safe distance, stopping, and asking for a path when the path was completely blocked-in different behavior conditions. As an outcome of participant responses, we found that participants felt more comfortable in the third behavior condition among all the conditions in terms of perceived intelligence, likeability, and perception. Further, participants perceived robots following a role of an assistant and also liked the inclusion of audio and visual aids in the robot in the third behavior condition. Hence, it can be connoted, that in a human-populated environment, a robot's behavior that exhibits social behavior is more acceptable.

In the future, we hope to equip the PI with a robust communication protocol, to live stream all, the activities. Further, we look forward to adapting optimization algorithms for navigation planning [37,38] and other deep learning algorithms for vision processing [39–46] to test the PI in various other social spaces like malls, restaurants, etc.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] K. A. Demir, G. Döven and B. Sezen, "Industry 5.0 and human-robot co-working," *Procedia Computer Science*, vol. 158, no. 1, pp. 688–695, 2019.

[2] World Robotics, "*Executive Summary World Robotics 2017 Service Robots.* Frankfurt," *Germany: International Federation of Robotics*, pp. 15–17, 2017. [Online]. Available: https://ifr.org/free-downloads/.

[3] World Robotics, "*Executive Summary World Robotics 2018 Service Robots.* Frankfurt," *Germany: International Federation of Robotics*, pp. 13–24, 2018. [Online]. Available: https://ifr.org/free-downloads/.

[4] S. Ivanov and C. Webster, "Perceived appropriateness and intention to use service robots in tourism," in *Information and Communication Technologies in Tourism*, Springer, Nicosia, Cyprus, 2019, pp. 237–248.

[5] S. Ivanov, C. Webster and A. Garenko, "Young Russian adults' attitudes towards the potential use of robots in hotels," *Technology in Society*, vol. 55, no. 2, pp. 24–32, 2018.

[6] S. Ivanov, C. Webster and P. Seyyedi, "Consumers' attitudes towards the introduction of robots in accommodation establishments," *Tourism*, vol. 66, no. 3, pp. 302–317, 2018.

[7] K. J. Singh, D. S. Kapoor and B. S. Sohi, "Selecting social robot by understanding human–robot interaction," in *Int. Conf. on Innovative Computing and Communications*, New Delhi, India, 2020, pp. 203–213.

[8] E. Leon, K. Amaro, F. Bergner, I. Dianov and G. Cheng, "Integration of robotic technologies for rapidly deployable robots," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1691–1700, 2017.

[9] G. Sawadwuthikul, T. Tothong, T. Lodkaew, P. Soisudarat, S. Nutanong *et al.,* "Visual goal human-robot communication framework with Few-shot learning: A case study in robot waiter system," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 13–25, 2021.

[10] H. Hong and Y. Yin, "Ontology-based conceptual design for ultra-precision hydrostatic guideways with human–Machine interaction," *Journal of Industrial Information Integration*, vol. 2, no. 3, pp. 11–18, 2016.

[11] X. Truong, V. Yoong and T. Ngo, "Socially aware robot navigation system in human interactive environments," *Intelligent Service Robotics*, vol. 10, no. 4, pp. 287–295, 2017.

[12] A. Vega, L. Manso, D. Macharet, P. Bustos and P. Núñez, "Socially aware robot navigation system in human-populated and interactive environments based on an adaptive spatial density function and space affordances," *Pattern Recognition Letters*, vol. 118, no. 3, pp. 72–84, 2019.

[13] D. Thang, L. Nguyen, P. Dung, T. Khoa, N. Son *et al.,* "Deep learning-based multiple objects detection and tracking system for socially aware mobile robot navigation framework," in *2018 5th NAFOSTED Conf. on Information and Computer Science (NICS)*, Ho Chi Minh City, Vietnam, 2018, pp. 436–441.

[14] R. Mead and M. J. Matarić, "Autonomous human–robot proxemics: Socially aware navigation based on interaction potential," *Autonomous Robots*, vol. 41, no. 5, pp. 1189–1201, 2017.

[15] Y. Lu, "Industry 4.0: A survey on technologies, applications and open research issues," *Journal of Industrial Information Integration*, vol. 6, no. 2, pp. 1–10, 2017.

[16] N. Motoi, M. Ikebe and K. Ohnishi, "Real-time gait planning for pushing motion of humanoid robot," *IEEE Transactions on Industrial Informatics*, vol. 3, no. 2, pp. 154–163, 2007.

[17] P. Singh, N. Pisipati, P. Krishna and M. Prasad, "Social signal processing for evaluating conversations using emotion analysis and sentiment detection," in *2019 Second Int. Conf. on Advanced Computational and Communication Paradigms (ICACCP)*, Gangtok, Sikkim, 2019, pp. 1–5.

[18] C. Bartneck, D. Kulić, E. Croft and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[19] C. Mavrogiannis, A. M. Hutchinson, J. Macdonald, P. Alves-Oliveira and R. A. Knepper, "Effects of distinct robot navigation strategies on human behavior in a crowded environment," in *2019 14th ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, Daegu, South Korea, 2019, pp. 421–430.

[20] C. Vassallo, A. Olivier, P. Souères, A. Crétual, O. Stasse *et al.,* "How do walkers behave when crossing the way of a mobile robot that replicates human interaction rules?," *Gait Posture*, vol. 60, no. 2, pp. 188–193, 2018.

[21] R. Burns, M. Jeon and C. H. Park, "Robotic motion learning framework to promote social engagement," *Applied Sciences*, vol. 8, no. 2, pp. 241–253, 2018.

[22] A. Rossi, F. Garcia, A. C. Maya, K. Dautenhahn, K. L. Koay *et al.,* "Investigating the effects of social interactive behaviours of a robot on people's trust during a navigation task," in *Annual Conf. Towards Autonomous Robotic Systems*, London, UK, 2019, pp. 349–361.

[23] S. Warta, O. Newton, J. Song, A. Best and S. Fiore, "Effects of social cues on social signals in human-robot interaction during a hallway navigation task," *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 62, no. 1, pp. 1128–1132, 2018.

[24] S. Chatterjee, F. Zunjani and G. Nandi, "Real-time object detection and recognition on low-compute humanoid robots using deep learning," in *2020 6th Int. Conf. on Control, Automation and Robotics (ICCAR)*, Singapore, 2020, pp. 202–208.

[25] Z. Luo, A. Small, L. Dugan and S. Lane, "Cloud chaser: Real time deep learning computer vision on low computing power devices," in *Eleventh Int. Conf. on Machine Vision (ICMV 2018)*, Munich, Germany, 2018, pp. 743–750.

[26] D. Reis, D. Welfer, M. Cuadros and D. Gamarra, "Mobile robot navigation using an object recognition software with RGBD images and the YOLO algorithm," *Applied Artificial Intelligence*, vol. 33, no. 14, pp. 1290–1305, 2019.

[27] M. Graaf and B. Malle, "People's judgments of human and robot behaviors: A robust set of behaviors and some discrepancies," in *Companion of the 2018 ACM/IEEE Int. Conf. on Human-Robot Interaction*, Chicago USA, 2018, pp. 97–98.

[28] K. J. Singh, A. Nayyar, D. S. Kapoor, N. Mittal, S. Mahajan *et al.,* "Adaptive flower pollination algorithm-based energy efficient routing protocol for multi-robot systems," *IEEE Access*, vol. 9, pp. 82417–82434, 2021.

[29] P. Hancock, D. Billings, K. Schaefer, J. Chen, E. Visser *et al.,* "A Meta-analysis of factors affecting trust in human-robot interaction," *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011.

[30] M. Walters, K. Dautenhahn, K. Koay, C. Kaouri, R. Boekhorst *et al.,* "Close encounters: Spatial distances between people and a robot of mechanistic appearance," in *5th IEEE-RAS Int. Conf. on Humanoid Robots*, Tsukuba, Japan, 2005, pp. 450–455.

[31] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Las Vegas, Nevada, 2016, pp. 779–788.

[32] S. Hsu, Y. Wang and C. Huang, "Human object identification for human-robot interaction by using fast R-CNN," in *2018 Second IEEE Int. Conf. on Robotic Computing (IRC)*, CA, USA, 2018, pp. 201–204.

[33] A. C. Rios, D. H. Reis, R. M. Silva, M. A. Cuadros and D. F. T. Gamarra, "Comparison of the YOLOv3 and SSD mobileNet v2 algorithms for identifying objects in images from an indoor robotics dataset," in *2021 14th IEEE Int. Conf. on Industry Applications (INDUSCON)*, Sao Paulo, Brazil, 2021, pp. 96–101.

[34] S. Kert, M. Erkoç and S. Yeni, "The effect of robotics on six graders' academic achievement, computational thinking skills and conceptual knowledge levels," *Thinking Skills and Creativity*, vol. 38, no. 3, pp. 100714, 2020.

[35] C. Strabmann, S. Eimler, A. Arntz, A. Grewe, C. Kowalczyk *et al.,* "Receiving robot's advice: Does it matter when and for what?," in *Int. Conf. on Social Robotics*, Golden, USA, 2020, pp. 271–283.

[36] R. Wullenkord and F. Eyssel, "The influence of robot number on robot group perception - A call for action," *ACM Transactions on Human-Robot Interaction*, vol. 9, no. 4, pp. 1–14, 2020.

[37] H. Singh, M. Abouhawwash, N. Mittal, R. Salgotra, S. Mahajan *et al.,* "Performance evaluation of non-uniform circular antenna array using integrated harmony search with differential evolution based naked mole rat algorithm," *Expert Systems with Applications*, vol. 189, no. 1, pp. 116146, 2022.

[38] R. Salgotra, M. Abouhawwash, U. Singh, S. Saha, N. Mittal *et al.,* "Multi-population and dynamic-iterative cuckoo search algorithm for linear antenna array synthesis," *Applied Soft Computing*, vol. 113, no. 2, pp. 108004, 2021.

[39] S. Mahajan, A. Raina, M. Abouhawwash, X. Z. Gao and A. Pandit, "Covid-19 detection from chest X-ray images using advanced deep learning techniques," *Computers, Materials & Continua*, vol. 70, no. 1, pp. 1541–1556, 2022.

[40] T. Jabeen, I. Jabeen, H. Ashraf, N. Jhanjhi, H. Mamoona *et al.,* "A sultan a monte carlo based COVID-19 detection framework for smart healthcare," *Computers, Materials, & Continua*, vol. 70, no. 2, pp. 2365–2380, 2021.

[41] M. Abdel-Basset, N. Moustafa, R. Mohamed, O. Elkomy and M. Abouhawwash, "Multi-objective task scheduling approach for fog computing," *IEEE Access*, vol. 9, no. 14, pp. 126988–127009, 2021.

[42] M. Abouhawwash and A. Alessio, "Multi-objective evolutionary algorithm for PET image reconstruction: Concept," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 1–10, 2021.

[43] L. Gaur, U. Bhatia, N. Z. Jhanjhi, G. Muhammad and M. Masud, "Medical image-based detection of COVID-19 using deep convolution neural networks," *Multimedia Systems*, vol. 3, no. 4, pp. 1–10, 2021.

[44] A. Tayal, J. Gupta, A. Solanki, K. Bisht, A. Nayyar *et al.,* "DL-CNN-based approach with image processing techniques for diagnosis of retinal diseases," *Multimedia Systems*, vol. 12, no. 3, pp. 1–22, 2021.

[45] M. Masud, G. S. Gaba, K. Choudhary, R. Alroobaea and M. S. Hossain, "A robust and lightweight secure access scheme for cloud based E-healthcare services," *Peer-to-peer Networking and Applications*, vol. 14, no. 3, pp. 3043–3057, 2021.

[46] M. Masud, G. Gaba, S. Alqahtani, G. Muhammad, B. Gupta *et al.,* "A lightweight and robust secure key establishment protocol for internet of medical things in COVID-19 patients care," *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15694–15703, 2021.