Perceived Appropriateness: A Novel View for Resolving Inappropriate Robot Behavior in Socially-Aware Navigation

Zhou, Y.

DOI
10.4233/uuid:2b8953d2-e29a-4185-9dbd-b99548d10365

Publication date
2025

Document Version
Final published version

Citation (APA)
Zhou, Y. (2025). *Perceived Appropriateness: A Novel View for Resolving Inappropriate Robot Behavior in Socially-Aware Navigation*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:2b8953d2-e29a-4185-9dbd-b99548d10365

Perceived Appropriateness: A Novel View for Resolving Inappropriate Robot Behavior in Socially-Aware Navigation

Perceived Appropriateness: A Novel View for Resolving Inappropriate Robot Behavior in Socially-Aware Navigation

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr. ir. T. H. J. J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
March 20th 2025 at 10 o'clock

by

Yunzhong ZHOU

Master of Engineering in Mechanical Engineering
Tongji University, China
born in Zhejiang Province, China

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof. dr. G. Kortuem,	Delft University of Technology, Promotor
Prof. dr. ir. A. Bozzon,	Delft University of Technology, Promotor
Independent members:	
Prof. dr. C.M. Jonker	Delft University of Technology
Prof. dr. A. Spalanzani	Institut national de recherche en informatique, France
Prof. dr. J. Finney	Lancaster University, United Kingdom
Prof. dr. M.H. Martens	Eindhoven University of Technology
Reserve member:	
Prof. dr. ir. K.M.B. Jansen	Delft University of Technology

Keywords: Socially Aware Navigation, Mobile Robot, Social Signal Processing, Perceived Appropriateness, Human-Robot Interactions

Printed by: RIDDERPRINT || www.ridderprint.nl/en/

Cover by: Yunzhong Zhou

ISBN/EAN: 978-94-6518-029-8

An electronic copy of this dissertation is available at
https://repository.tudelft.nl/.

Contents

Summary

Robots are increasingly navigating our living environments and must navigate socially to be accepted. While existing socially aware navigation (SAN) approaches enable robots to interpret and communicate social information to navigate efficiently, safely, and in a socially acceptable manner, they often overlook potential conflicts and errors in real-world human-robot interactions. This thesis contributes to SAN by investigating how robots can adapt their inappropriate navigation behavior based on human feedback (perceived appropriateness), leading to smoother and less error-prone human-robot interactions.

At the core of SAN approaches is the understanding and processing of information communicated between humans and robots, enabling agents to adapt to each other. Despite extensive research, most studies focus on specific aspects of human-robot interactions, limiting a comprehensive understanding of robot information communication capabilities. In Chapter 2, we conducted a scoping review to synthesize the information investigated in the existing literature to identify research gaps and inform future investigations. The information was categorized into future behaviors, social signals, and proxemic zones, where social signals were the most informative as they influence other information. Within the social signals examined, the information of perceived appropriateness (PA) of robot navigation behavior was the most informative. It directly indicated the appropriateness of robot behavior, facilitating informed adaptations directly related to human preferences and needs. However, this crucial information has remained the least investigated and deserves further investigation.

A field observation study was conducted in Chapter 3 to investigate human PA and responses to robot navigation behaviors in conflicting situations. The results revealed that narrow spaces lead to richer perceived appropriateness and higher expectations for robot adaptations. The study identified environmental factors (e.g., path width), human demographics, and robot behaviors as significant contributors to human yielding behavior and PA. This underscores the importance of studying PA in narrow environments.

Leveraging prior insights, the PASNiP dataset for PA detection was created via human-robot interaction experiments featuring designed robot behaviors in narrow environments to capture a complete range of PA levels, in Chapter 4. This dataset includes extensive cues for PA detection, including human and robot motion features, as well as human emotion and attention. Machine learning models applied to the dataset demonstrate that incorporating emotional and attentional features significantly improves PA detection accuracy, from 63% to 68% (algorithm-predicted) and 79% (participant-reported). This suggests that emotions, attention, and potentially other social signals are effective for PA detection.

Assuming that human PA regarding robot navigation behavior could be detected, Chapter 5 further explored how humans prefer robots to adapt their inappropriate navigation behaviors in narrow spaces. The findings revealed a strong preference for robots to move aside and stop, regardless of the types of inappropriate behavior. This preference contrasts with findings in non-navigation settings, likely due to the increased safety concerns associated with inappropriate navigation behaviors. This suggests a generic adaptation strategy for robots in narrow spaces and highlights the need for further research into context-specific adaptations.

Overall, this thesis identifies PA as a crucial social signal and has conducted systematic investigations to enable robots to understand and adapt their inappropriate behaviors/errors in real-time, ultimately enhancing their social acceptance.

Samenvatting

Robots navigeren steeds vaker in onze leefomgevingen en moeten sociaal navigeren om geaccepteerd te worden. Hoewel bestaande sociaal bewuste navigatie (SAN) benaderingen robots in staat stellen sociale informatie te interpreteren en te communiceren om efficiënt, veilig en op een sociaal aanvaardbare manier te navigeren, negeren ze vaak potentiële conflicten en fouten in real-world interacties tussen mens en robot. Dit proefschrift draagt bij aan SAN door te onderzoeken hoe robots hun ongepast navigatiegedrag kunnen aanpassen op basis van menselijke feedback (waargenomen gepastheid), wat leidt tot soepelere en minder foutgevoelige interacties tussen mens en robot.

De kern van SAN-benaderingen is het begrijpen en verwerken van informatie die tussen mensen en robots wordt gecommuniceerd, waardoor agenten zich aan elkaar kunnen aanpassen. Ondanks uitgebreid onderzoek, richten de meeste studies zich op specifieke aspecten van interacties tussen mens en robot, waardoor een alomvattend begrip van de mogelijkheden van robots voor informatiecommunicatie wordt beperkt. In hoofdstuk 2 hebben we een scoping review uitgevoerd om informatie die in bestaande literatuur is onderzocht te synthetiseren, om hiaten in het onderzoek te identificeren en toekomstige onderzoeken te informeren. Hoofdstuk 2 categoriseerde de informatie in toekomstig gedrag, sociale signalen en proxemische zones, waarbij sociale signalen het meest informatief waren omdat ze andere informatie beïnvloeden. Binnen de onderzochte sociale signalen was de informatie over de waargenomen gepastheid (PA) van het navigatiegedrag van de robot het meest informatief. Het gaf direct de gepastheid van het robotgedrag aan, waardoor weloverwogen aanpassingen direct gerelateerd aan menselijke voorkeuren en behoeften mogelijk werden gemaakt. Deze cruciale informatie bleef echter het minst onderzocht en verdient nader onderzoek.

Een veldobservatiestudie werd uitgevoerd in hoofdstuk 3 om de menselijke PA van en reacties op het navigatiegedrag van robots in conflictsituaties te onderzoeken. De resultaten toonden aan dat smalle ruimtes leiden tot een rijkere waargenomen gepastheid en hogere verwachtingen voor robotaanpassingen. De studie identificeerde omgevingsfactoren (bijv. padbreedte), menselijke demografie en robotgedrag als belangrijke factoren die bijdragen aan menselijk voorrangsgedrag en PA. Dit onderstreept het belang van het bestuderen van PA in smalle omgevingen en benadrukt de rol van sociale signalen, aandacht en emotie bij het detecteren van PA.

Gebruikmakend van eerdere inzichten, werd de PASNiP-dataset voor PA-detectie gecreëerd via experimenten met interactie tussen mens en robot, met ontworpen robotgedragingen in smalle omgevingen om een compleet scala aan PA-niveaus vast te leggen 4. Deze dataset bevat uitgebreide signalen voor PA-detectie, waaronder bewegingskenmerken van mens en robot, evenals menselijke emotie en aandacht.

Machine learning-modellen die op de dataset zijn toegepast, tonen aan dat het opnemen van emotionele en aandachtsgerichte kenmerken de nauwkeurigheid van PA-detectie aanzienlijk verbetert, van 63% naar 68% (algoritme-voorspeld) en 79% (door deelnemers gerapporteerd). Dit suggereert dat emoties, aandacht en mogelijk andere sociale signalen effectief zijn voor PA-detectie.

Gebaseerd op de aanname dat de menselijke PA van het navigatiegedrag van de robot kan worden gedetecteerd, onderzocht hoofdstuk 5 verder hoe mensen willen dat robots hun ongepast navigatiegedrag in smalle ruimtes aanpassen. De bevindingen lieten een sterke voorkeur zien voor robots om aan de kant te gaan en te stoppen, ongeacht het type ongepast gedrag. Deze voorkeur contrasteert met bevindingen in niet-navigatieomgevingen, waarschijnlijk vanwege de verhoogde veiligheidsrisico's die gepaard gaan met ongepast navigatiegedrag. Dit suggereert een generieke aanpassingsstrategie voor robots in smalle ruimtes en benadrukt de noodzaak van verder onderzoek naar contextspecifieke aanpassingen.

Over het geheel genomen identificeert dit proefschrift PA als een cruciaal sociaal signaal, en zijn er systematische onderzoeken uitgevoerd om robots in staat te stellen hun ongepast gedrag/fouten in realtime te begrijpen en aan te passen, wat uiteindelijk hun sociale acceptatie bevordert.

1

INTRODUCTION

> “It takes humility to seek feedback. It takes wisdom to understand it, analyze it, and appropriately act on it.” - Stephen Covey

Robots are increasingly navigating into social environments, taking on roles that promise significant benefits. In logistics, robots significantly reduce costs by automating tasks such as package sorting, loading, and delivering [1]. In healthcare, robots improve operational workflows and enhance patient comfort by aiding with mobility and offering companionship to the elderly [2]. In domestic settings, robotic vacuum cleaners and personal assistants handle mundane household tasks, thereby liberating time for homeowners [3]. Fundamental to their successful operations in social environments—whether delivering packages on busy urban streets, assisting patients in hospitals, or performing household tasks—is their ability to navigate effectively around humans. According to the computers-as-social-actors theory [4, 5], robots are perceived as social agents and expected to behave socially. This has led to burgeoning studies of socially aware navigation (SAN), also known as human-aware navigation [6], socially compliant navigation [7], socially acceptable navigation [8], and socially competent navigation [9], as reviewed in Chapter 1.1. Despite advancements in SAN, robots often operate under the assumption of perfection, yet they are prone to occasional errors in real-world applications, potentially undermining human trust and acceptance. In extreme cases, such errors could result in restrictions or bans on their use in social environments. An exception is the study by Vroon et al. [10], which introduced the concept of “perceived appropriateness (PA)" to describe robot errors during navigation. This concept refers to the human subjective perception of robot behavior appropriateness, as whether a robot makes an error or not depends solely on human perceptions, as discussed in Chapter 1.2. In contrast, many studies have investigated robot errors in the field of human-robot interactions (HRI), which collectively contribute to both interaction quality and robot acceptance, as reviewed in Chapter 1.3. Inspired by these studies in HRI, we propose approaches to investigate robot errors/PA within socially aware navigation, further detailed in Chapter 1.4. Building on these approaches, we conduct a series of studies to complete the thesis, with the outline and key contributions presented in Chapter 1.5.

1.1. BACKGROUND: SOCIALLY AWARE NAVIGATION (SAN)

The field of socially aware navigation (SAN) has witnessed a notable increase in academic studies, particularly following 2010, as depicted in Figure 1.1 [11]. This surge highlights the growing interest in enhancing robotic navigation capabilities within social contexts.

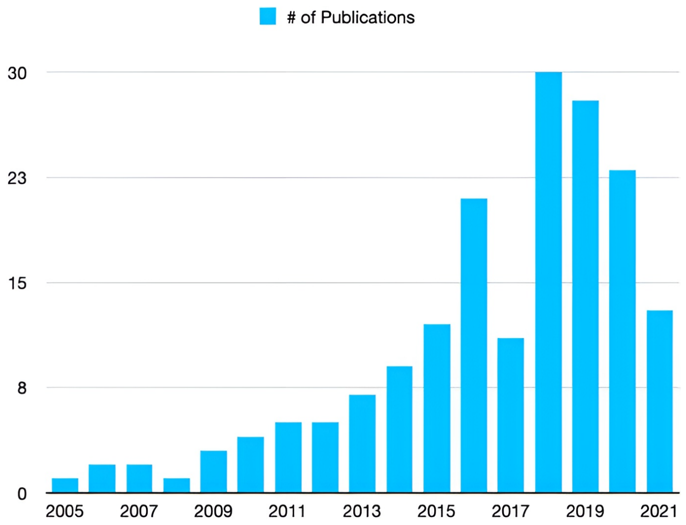


Figure 1.1: Number of published SAN studies over the years.

A socially aware robot navigation framework typically comprises two essential components: a traditional navigation module and a social awareness module [12]. The social awareness module allows robots to go beyond conventional navigation strategies by incorporating social conventions and computational methods. By doing so, robots can navigate more efficiently in human environments, adjusting their behavior to align with social norms and expectations.

1.1.1. SOCIAL CONVENTIONS/PRINCIPLES

Agents often find it challenging to understand exactly what other agents, particularly humans, aim to accomplish, how they feel, or what they prefer. Additionally, social norms that could offer guidance are frequently unspoken. Therefore, researchers have identified many principles of social navigation, as summarized by a recent review into 8 categories including *(1) safety, (2) comfort, (3) legibility, (4) politeness, (5) social competency, (6) understanding other agents, (7) proactivity, and (8) responding appropriately to context*, as visualized in Figure 1.2. These principles have been widely applied in socially aware navigation studies. For example, the principle of safety can be ensured by avoiding conflicts [13] or respecting human personal or intimate zones [14–16]. The principle of legibility can be achieved by optimizing robot paths or communicating intent through speech or light signals [17, 18]. The principle of politeness can be reached by the robot giving priority to humans and displaying low dominance paths [19]. These principles collectively contribute to the increasingly successful applications of robots into social environments.

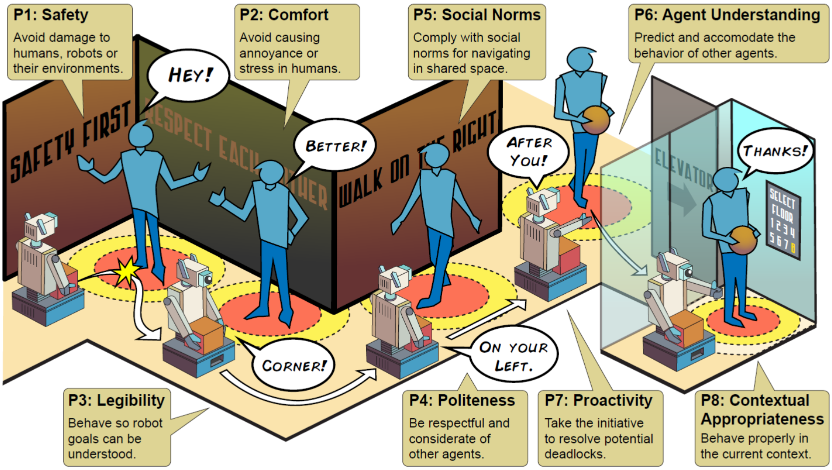


Figure 1.2: Principles guiding robot behavior design

1.1.2. COMPUTATIONAL METHODS

Enabling robots to adhere to social conventions/principles fundamentally requires processing social cues, which are salient and transient features, allowing others to pick up useful information [20]. Several methods have been proposed to achieve this, including behavior prediction, the semantics of/proxemic spaces, and social signal processing.

Behavior prediction Robots can enhance safety, comfort, and social interactions by predicting human trajectories and (re-)planning paths, thereby avoiding collisions and conflicts [13]. This prediction relies on datasets of human walking trajectories [21], considering factors like neighboring pedestrians [22], scene layouts [23], and attention dynamics [24], particularly during explicit interactions [25]. Additional factors, including group cohesion and directional tendencies [26], as well as interactions between humans and agents or objects [27, 28], further contribute to behavior prediction performances. Robots could also exhibit more predictable behaviors to humans, known as legibility, through movements and supplementary cues such as gaze, light signals, and sound [17, 29–37].

Proxemic space Proxemics investigates the spatial boundaries people establish in diverse social and relational contexts according to Hall [38]. The proxemics space can be classified into four zones: intimate (≤ 0.45 m), personal (> 0.45 m), social (> 1.2 m), and public (> 3.6 m). These zones are largely modelled through asymmetric

Gaussian functions, considering human position, walking speed, direction, emotion, and dominance [39–43]. To ensure safe and effective interactions, robots are programmed to avoid or minimize intrusion into the personal or intimate zones [44–46] and initiate interactions within social zones.

Social Signal Processing Employing social signal processing (SSP) approaches enables robots to interpret human social signals and respond effectively, including emotion, intention, and dominance [47, 48]. For example, robots control their distance from humans based on detected emotional states [49–51], or alter their paths based on human dominance levels [52, 53]. Robots also utilize SSP to help humans better understand their behaviors, such as using facial expressions and movement patterns to convey emotions and social presence [54–57].

1.2. LIMITATIONS: FEW STUDIES ON ERRORS/PA IN SAN

Current socially aware navigation approaches have predominantly focused on enhancing robots' social abilities, often operating under the assumption that robots function flawlessly. In real-life social navigation scenarios, robots are still prone to various errors, both in terms of hardware limitations and interaction dynamics. The focus on perfect interaction may cause researchers to overlook crucial aspects that could undermine interaction quality.

What constitutes a robot error or mistake? In the context of socially aware navigation, errors are typically only acknowledged when they result in a collision. However, existing socially aware navigation techniques—particularly proxemic space modelling and behaviour prediction—have already enabled robots to maintain safe distances and avoid moving towards humans, thus preventing collisions or conflicts [58]. Even when navigating in crowded environments, robots could still stop for extended periods to prevent collisions, a challenge referred to as the "freezing-robot problem (FRP)". Of course, many studies have also tried to address the FRP, thus preventing robots from erroneous situations [59, 60]. In sum, few studies directly examine instances of what was traditionally defined as robot errors and explore the dynamics of these interactions. This gap may stem from the limited understanding of what constitutes an error in the context of social navigation.

In socially aware navigation, whether a robot makes an error depends solely on the subjective perception of humans. As such, Jered et al. [10] introduced the concept of "perceived appropriateness (PA)", which refers to the human subjective perception of the appropriateness of robot behavior. The study collected a dataset and enabled the detection of PA in robot positioning behavior. However, it only identified inappropriate robot positioning and did not explore how robots could adapt their behaviors to resolve these errors. The current limited understanding of errors/PA in robot social navigation hinders the ability of robots to correct these errors effectively, thereby reducing the smoothness of their interactions and responsiveness to individual preferences and needs. In contrast, extensive error studies have already been conducted in the field of human-robot interactions, which has significantly contributed to enhancing the overall functionality of robots.

1.3. INSIGHTS: ERROR STUDIES IN HUMAN-ROBOT INTERACTION

Over the decades, researchers have extensively explored error detection and resolution in human-robot interactions, commonly defining such incidents also as “failures", “mistakes", and “faults". These studies have focused on different aspects of error situations as visualized in Figure 1.3, including factors underlying error occurrences, error detection/prediction, and error repair/resolution [61].

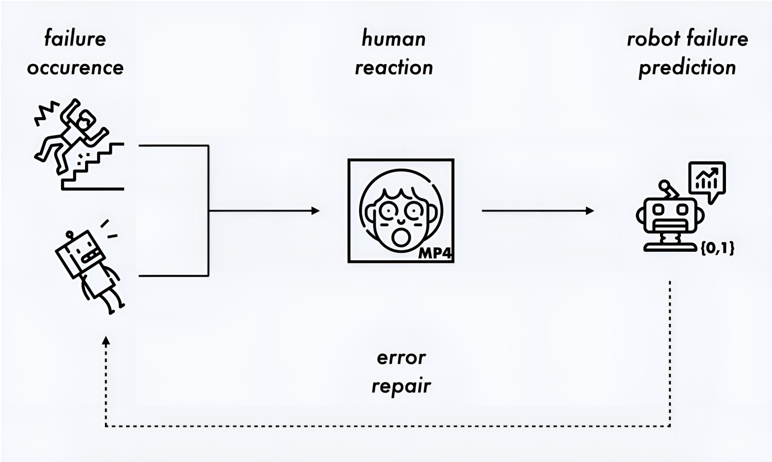


Figure 1.3: Overview of error/failure detection and repair

Factors underlying error occurrences Numerous studies have sought to identify and analyze error instances and the underlying factors contributing to their occurrence. Errors are typically categorized into technical failures (from mechanical or system deficiencies) and social norm violations (when robots deviate from expected social behavior) [62]. Technical failures are typically caused by factors related to software bugs or hardware. However, the factors driving social norm violations are context-dependent. In human-robot conversations, errors often stem from improper gaze coordination, speech interruptions, or prolonged silences, which disrupt natural communication patterns [63]. In cooperative Lego tasks, social norm violations arise from providing unclear task instructions or introducing communication delays that hinder task progression [64].

Human responses and error detection Robots inevitably make errors, particularly in the dynamics of human-robot interactions. Numerous studies have explored the impact of these errors on humans, revealing that they can significantly affect how humans perceive and respond to robots. Psychologically, research shows that errors reduce the perceived trustworthiness and reliability of robots [65]. Interestingly, humans sometimes find error robots more likable than those that operate flawlessly

1

[64], yet the task performance was significantly lower in the faulty robot condition [66]. Physically, humans demonstrate diverse reactions to robot errors. In human-robot cooperative tasks, participants responded to robot mistakes using a range of social cues, including speech, smiling or laughing, facial expressions, head and body movements, and gaze shifts [64, 67]. Another study revealed that about 80% of error occurrences resulted in visible human responses within an average of 3 seconds. Humans' physical responses, as revealed by social cues, have been widely used for detecting robot errors. These studies have utilized different social cues to enable error detection, including head and shoulder movements [68], facial expressions [69], or combining poses, facial, and audio features [69, 70] that result in an accuracy of about 70% or even higher. Additionally, the more critical the error and the greater the threat to the interaction flow, the more social cues people exhibit in response [63].

Error repair/resolution Although errors decrease robot trustworthiness and reliability, effective recovery strategies substantially enhance robot evaluation and overall interaction quality [63, 71]. Error studies in human-robot interactions reveal human preferences for error-specific repair strategies [63, 72]. Therefore, many studies have explored different robot errors and compared how varied repair strategies perform under these errors. For example, in a robot control system, different repair/recovery strategies have been used for different errors: When mode confusion occurs, the robot informs the user of necessary actions or triggers a reset; When a module hangup occurs, the robot asks the user to wait or inform the breakdown; When sensors detect an obstacle, the robot informs the user of the issue and requests assistance, such as asking the user to push it away manually [72]. In another collaborative task, it is revealed that the main effect of error type is significant on most of the Godspeed rating sub-scales averaged across all robot recovery strategies. Specifically, the robot recovery strategies are rated significantly higher concerning anthropomorphism if applied to planning errors (PE) compared to execution errors (EE) and social norm violations (SNV). The average likeability and perceived intelligence ratings of robot reactions are significantly higher if paired with PE and EE compared to SNV [63].

1.4. RESEARCH APPROACH: PA IN SAN

As reviewed in Chapter 1.2, the study of errors/PA in socially aware navigation is important but limited. In contrast, errors in human-robot interactions have been thoroughly examined, with three major aspects collectively contributing to better error management, as discussed in Chapter 1.3. Understanding errors and the underlying factors enables avoiding certain errors. Understanding and incorporating human responses (through social cues) into machine learning enhances robots' error awareness. Identifying error types and their relations with different error-repairing strategies improves human perception and interaction quality despite errors. Inspired by the error studies in HRI, this study aims to understand the underlying factors, human responses to, and the handling of different robot errors/PA in a socially aware navigation context. We propose the question: **When navigating social environments, how can robots understand and adapt to humans to enhance the perceived appropriateness of their behaviors?** We divide this overarching question into four specific research sub-questions. The core of this thesis is structured around these sub-questions, as detailed below:

1. Mobile robots often struggle to understand social environments and dynamic interaction [73, 74]. Fundamental to such social dynamics are social cues, which are salient features for communicating useful information. Therefore, we propose the research question:

RQ1: In existing research on socially aware navigation, what types of information communicated between humans and robots have been studied, and what social cues have been used for this communication?

2. Robots with navigation strategies (e.g., [75]), especially the social force model and its variations [76–78], occasionally make errors when pedestrians do not yield as anticipated and/or the yielding behaviors of robots are not adequate. As they are what triggers and necessitates yielding behaviors, we focus on **conflicts**, where a collision occurs unless either the robot or human changes direction or speed [79]. We investigate:

RQ2: What are the key factors contributing to human responses to conflicts in public spaces, and how do they influence human PA of robot navigation behaviours?

3. Like humans, who recognize and adapt to feedback, robots should detect the perceived appropriateness of their behavior by using social cues. However, current robots cannot detect PA beyond mere positioning behavior [10]. Yet, such PA or conflicts largely happen in narrow spaces where agents struggle to navigate [80]. Therefore, we propose to address:

RQ3: How can we detect the perceived appropriateness of robot navigation behavior in narrow environments?

4. Based on detected PA, robots should further adapt their inappropriate behavior to mitigate humans' negative perceptions and enhance their social acceptance [64]. We aim to investigate further:

RQ4: How do humans prefer a robot to adapt its inappropriate navigation behavior in narrow environments?

Addressing these sub-questions will provide comprehensive answers to the main research question of this thesis, thus enhancing human perceptions of robots, even in the presence of errors, as illustrated in Figure 1.4. Figure 1.4.a enumerates the research questions and outlines their objectives, while Figure 1.4.b offers a visual representation of potential shifts in human perceptions of robots during interactions, post-resolution of each research question. Initially, human perceptions may deteriorate when the robot lacks awareness of PA information, progressing to a more positive perception despite errors after appropriate adaptations.

This improvement in perception is expected to result from systematically addressing research questions RQ1 to RQ4. Specifically, exploring RQ1 involves studying the information communication within Socially Aware Navigation (SAN) to uncover the importance of errors or PA. Robots repeat their errors without PA information,

1

leading to a gradual decline in human perception. This motivates RQ2, which aims to analyze the factors contributing to these errors or PA and to identify essential human social cues for detecting PA. Understanding these factors will enable robots to avoid specific errors, thus reducing the adverse effects on human perceptions.

Nevertheless, expecting robots to operate free from errors/PA is unrealistic. Therefore, addressing RQ3 is vital as it enables robots to detect PA and thus enables generic adaptations, such as stopping or moving away, to reduce negative human perceptions further. Humans prefer robots that tailor their adaptive responses to specific error/PA types. Thus, exploring RQ4 could lead to error-specific adaptations, allowing robots to adapt more effectively and thus significantly enhancing human perceptions.

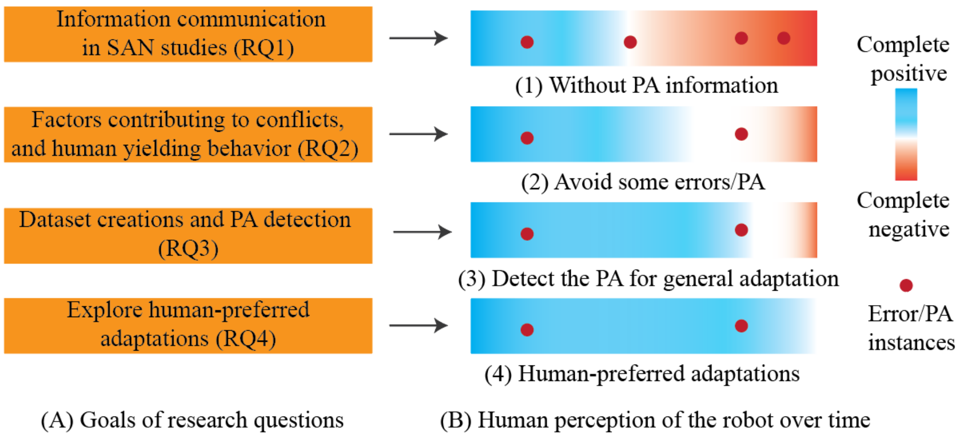


Figure 1.4: Expected human perception of a robot by addressing research questions

1.5. THESIS OUTLINE AND CONTRIBUTIONS

This dissertation is composed of six parts, including the Introduction (Chapter 1) and the Conclusion (Chapter 6). Each of Chapters 2, 3, 4, and 5 addresses one research question and is a peer-reviewed article. The thesis structure is shown in Figure 1.5.

1.5.1. PART I: PA IMPORTANCE

Chapter 2 addresses research question RQ1. This scoping review provides a nuanced understanding of human-robot information communication in navigation and identifies the important yet overlooked feedback information of perceived appropriateness (PA) [10], which deserves further exploration. This Chapter is under review as follows:

Shared Space, Shared Information: A Scoping Review of Social Cues and Information in Social Robot Navigation. Yunzhong Zhou, Jered Vroon, Zoltan Rusak, Gerd Kortuem. Under review in the International Journal of Social Robotics.

In the above publication, the author collected and analyzed relevant studies to understand information communication and wrote most of the publication's content. Other authors collectively analyzed a set of studies for inter-rater reliability.

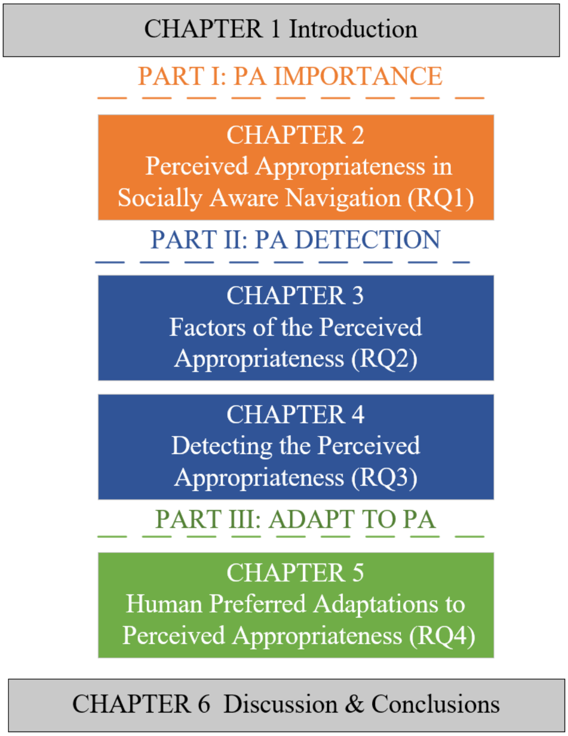


Figure 1.5: Thesis Outline

1.5.2. Part II: PA detection

Chapter 3 addresses research question RQ2. This Chapter investigates factors contributing to potential human-robot conflicts (important for triggering PA) and human-yielding behavior and investigate social cues involved in conflicting situations. This Chapter is under review as follows:

Streetwise Social Navigation: Field Observations of Yielding in Robot-Pedestrian Interactions. Jered Vroon, Yunzhong Zhou, Gerd Kortuem. ACM Transactions on Human-Robot Interaction (2024).

In the above publication, the author was involved in field observations and wrote some parts of the publication's content.

Chapter 4 addresses research question RQ3. We present a novel dataset, PARSNiP, for PA detection, which covers a complete PA range and includes emotion and attention. This Chapter is published as follows:

PARSNiP: A Novel Dataset for Better Perceived Appropriateness Detection in Robot Social Navigation using Emotion and Attention Features. Yunzhong Zhou, Jered Vroon, Zoltan Rusak, Gerd Kortuem. Under review in the International Journal of Social Robotics.

The author conducted the experiment in the above publication, collected and processed data to create the dataset, trained machine learning models for PA detection, and wrote most of the publication's content. The rest of the authors were collectively involved in the design of the experiment and supervised the research.

1.5.3. PART III: ADAPTING TO PA

Chapter 5 addresses research question RQ4. We collected 42 human-preferred robot adaptations to inappropriate robot navigation behavior and compared a selection of 12 for adapting to different PA types. This Chapter is publishedas follows:

This chapter has been published as: Zhou Y, Vroon J, Kortuem G. **Exploring Human Preferences for Adapting Inappropriate Robot Navigation Behaviors: A Mixed-Methods Study[J]**. IEEE Robotics and Automation Letters, 2024.

In the above publication, the author conducted the experiment, collected and analyzed data, and wrote most of the publication's content. The rest of the authors were collectively involved in the design of the experiment and supervised the research.

2

A SCOPING REVIEW OF INFORMATION IN SOCIALLY AWARE NAVIGATION

The successful integration of robots into our living environments fundamentally relies on their effective communication with humans. We conducted a scoping review to answer the research question: In existing research on socially aware navigation, what types of information communicated between humans and robots have been studied, and what social cues have been used for this communication (RQ1)? A systematic search on Scopus and Web of Science identified 11,300 studies, of which 176 matched our eligibility criteria. We analyzed the information investigated in these studies and categorized it into three clusters: future behaviors, social signals, and proxemic spaces. We further mapped social cues and their relations with different information, thus providing more systematic insights into human-robot communication. The results revealed that while most studies have focused on robots interpreting human behavior, there has been insufficient emphasis on human perceptions of robots, particularly real-time feedback such as perceived appropriateness, which is critical for adapting robot behavior to human preferences. Therefore, we advocate for further investigations into robots' capabilities for understanding, detecting, and correcting their inappropriate navigation behaviors, thus fostering their seamless integration into social environments.

2.1. INTRODUCTION

Robots are increasingly assuming societal roles such as delivery [81], eldercare [82], and cleaning [83], and their perception by humans as social actors necessitates their social intelligence [4, 5]. This has led to the development of socially-aware

This chapter has been submitted as: Yunzhong, Z., Vroon, J., Zoltan, Z., Gerd, K., 2024, Shared Space, Shared Information: A Scoping Review of Social Cues and Their Roles in Social Robot Navigation.

navigation. Unlike traditional approaches that consider human obstacles, socially aware navigation views humans as cooperative agents and emphasizes robot navigation that aligns with social norms and expectations.

Socially aware navigation relies fundamentally on effective human-robot communication [20]. Social signal processing (SSP) approaches have been widely developed to facilitate such communication [84]. SSP aims to provide computers with the ability to sense and understand human behavior by processing social cues [47, 48], which are channels of useful information. These cues can be clustered into five major categories, including physical appearance (height, attractiveness, body shape), gesture and posture (hand gestures, posture, walking), face and eyes behavior (facial expressions, gaze behavior, focus of attention), vocal behavior (prosody, turn taking, vocal outbursts), and space and environment (space and environment, distance, seating arrangement) [47]. Specific to navigation context, many studies have used SSP approaches to enable robots to safely and effectively navigate around humans [28], with more fine-grained social cues such as position, orientation, velocity, trajectory, motion, etc [12, 85]. These robots are not limited to detecting and responding to human social signals, such as emotions [42]; they could also predict human trajectories and plan their paths [58]. Furthermore, they could also model and respect personal space, vital for ensuring human safety and comfort [12]. Figure 2.1 shows a scene where a robot retreats by interpreting an approaching human's facial expressions and posture to respect the individual's negative emotion.

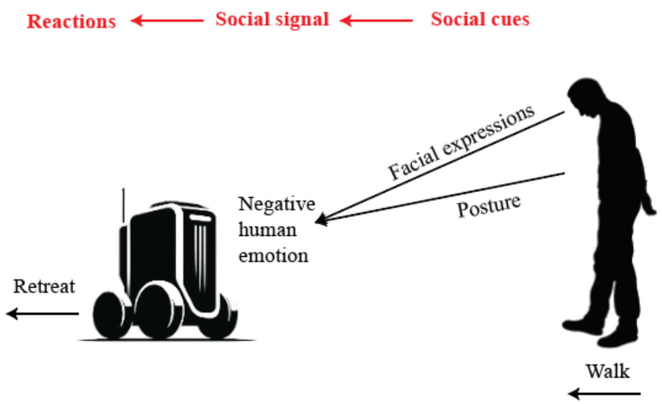


Figure 2.1: A robot processes human social cues for social navigation.

However, robots still struggle to understand social environments and dynamic interactions, triggering public concerns. For instance, mobile robots in San Francisco and at the University of Pittsburgh faced bans following numerous complaints about blocking pedestrians' paths [73, 74]. Recent literature reviews have investigated how robots can communicate and navigate effectively around humans. Rios-Martinez et al. explored the complexities of human-robot interactions, emphasizing the significance of robots adhering to human social norms such as proxemics and social signals [12]. Pascher et al. investigated how social cues communicate robots' movement intentions, highlighting the importance of social cues for social navigation [86]. Furthermore, Möller et al. reviewed various solutions for socially compliant navigation, stressing the importance of robots understanding and predicting human behavior [87].

Despite ongoing progress, we still lack a comprehensive understanding of the communication between humans and robots. Therefore, we seek to understand the types of information investigated in current studies and how they have been communicated through social cues by addressing the following research questions:

RQ1.1: What are the extrinsic characteristics of the studies in socially aware navigation?

RQ1.2: How did robot social navigation studies investigate human-robot information communication?

2.2. METHOD

2.2.1. LITERATURE SEARCH AND STUDY SELECTION

Our literature search was performed in January 2024 using two databases: Scopus and Web of Science. Keywords combined the social environment ("social*", "human", "people"), robotics ("robot", "aware"), and navigation context ("navigat*"): TITLE-ABS-KEY(("social*" OR "human" OR "people") AND ("robot" OR "aware") AND "navigat*"). This was limited to peer-reviewed articles from journals or conferences without any constraints regarding publication dates. Both automated and manual methods were employed to remove duplication, yielding a total of 11,300 unique studies. Article screening occurred in two stages (i.e., title and abstract screening and full article screening), which resulted in 176 studies, as shown in the PRISMA flow diagram (Fig. 2.2). At each stage, the first author conducted initial article screening, and the results were checked for reliability by the second author. Discrepancies regarding the evaluation of papers between the authors were resolved through discussion with other authors until a consensus was reached. Articles screened by title and abstract were excluded if they did not mention the robot in human/social environments if the robot was not navigating near humans (compared to the broader field of AVs or AUVs), if the robot was not an autonomous agent (compared to assistive tools), or if humans were not treated as social agents (as opposed to sole obstacles to avoid). During the full article screening, publications were excluded if they did not focus on the navigation behavior (compared to those solely focused on language processing), did not focus on communication (compared to SLAM techniques), did not involve human-robot interactions (compared to those solely predicting human trajectories), or did not involve the robot's decision-making (as opposed to designing the control interface or teleportation studies).

2.2.2. DATA EXTRACTION AND CODING

Despite extensive studies, the definition of social cues and their communicated information remains ambiguous. To define social cues and information, we first

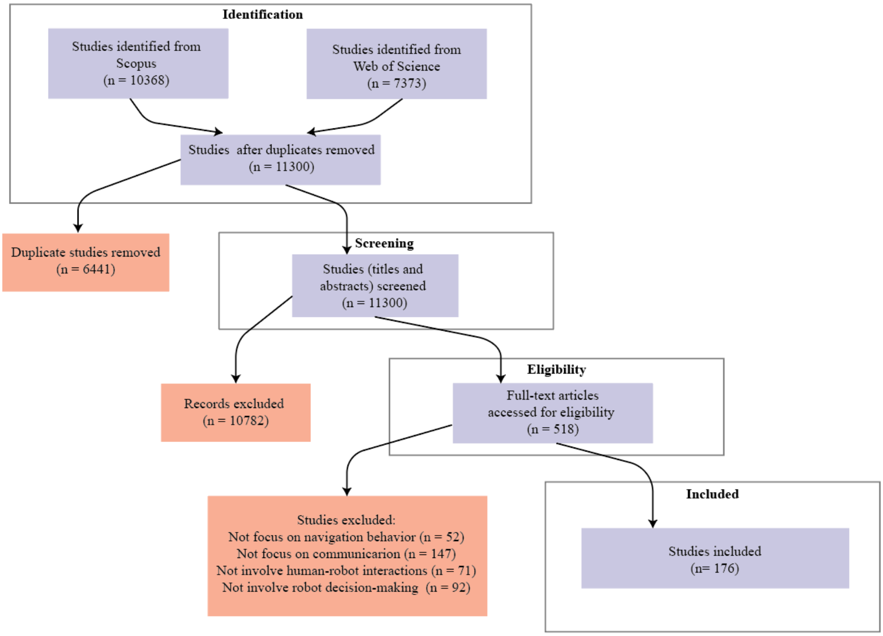


Figure 2.2: PRISMA flow diagram for selection of studies

determine an instance of human-robot communication as any data processing between a human and a robot involving input features and output data. In such communication, social cues are an agent's features serving as communication channels (capturing its communicative and agent-dependent nature [12, 20]), while information is the output data. Accordingly, the first author reviewed all studies to identify social cues and the information communicated between agents to design the extraction form. All social cues in the studies were included to ensure comprehensive coverage. These social cues not only included those widely recognized in human-robot interactions and robot social navigation (e.g., trajectory, gaze, posture) [12, 28] but also any other input features identified in the studies (e.g., position, external human-machine interfaces). Information was initially extracted using the original words mentioned in the studies. These include long-term predictions, short-term predictions, goals, personality, attitude, dominance, social presence, emotion, perceived appropriateness (PA), activity, group, personal space, personal zone, intimate space, intimate zone, social space, and (dynamic) social zone. A thematic analysis was then completed following the procedure by Braun and Clarke [88] to cluster similar information into larger categories. Based on group discussions and informed by the literature [9, 12, 89], we clustered information into three major categories: future behaviors (includes long-term predictions, short-term predictions), social signals (includes personality, attitude, dominance, social presence, emotion, perceived appropriateness (PA), activity, group), and proxemics zones (includes intimate, personal, social). Future behaviors are sequences composed of behavioral social cues/signals predicted [85]; social signals are expressions of one's attitude toward social situations and interplay [47]; proxemics are spatial distances individuals maintain in social and interpersonal situations [38].

We built a model to explain the information communication process in Figure 2.3. In this model, both the human and the robot alternately assume the roles of sender and receiver, thereby creating a dynamic interaction loop. Senders convey information using social cues such as gestures, speed, and positioning. At the same time, receivers interpret information (future behaviors, social signals, or proxemic zones) from these cues based on their models or past experiences. For example, a robot interprets a human's rapid approach as an intent to engage in urgent communication (the gray arrow towards the robot) and responds by moving towards the human to communicate its intention (the gray arrow towards the human). The robot also detects the human's negative emotions through facial expressions (the orange arrow towards the robot). It uses a soothing tone to communicate calm emotion (the orange arrow towards the human). Furthermore, by modeling the human's personal space (the green arrow towards the robot), the robot adjusts its position to maintain an appropriate distance. Humans and robots can achieve cooperative navigation by interpreting, responding to, and communicating information.

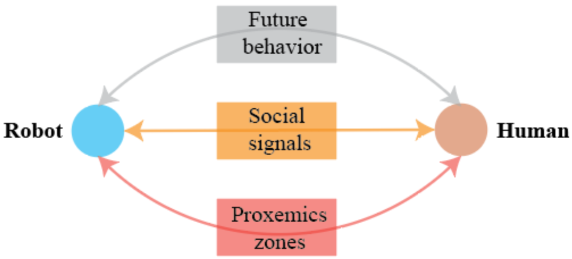


Figure 2.3: The communication model with different information communicated between humans and robots.

Having settled on definitions for social cues and information, we performed thorough coding and extraction. For each paper, we recorded:

- **General Data:** Author name, publication year, publication type (conference or journal), and publication channel.
- **Experimental Data:** The number of participants, age and gender distribution, the type of robot used, and the experimental setting (e.g., field or lab).
- **Communication Data:** The communicated information along with its category (e.g., emotion, social signal) and the social cues used for conveying this information (e.g., gesture, posture, trajectory).

Data extraction was initially performed by the first author of this article and subsequently reviewed by the second author. Differences were discussed and resolved once a consensus was reached.

2

2.3. THE EXTRINSIC CHARACTERISTICS OF STUDIES

Publication trend Figure 2.4 presents the year-wise publication trend. Between 1999 and 2010, 16 studies were published, contrasting with 160 studies published between 2011 and 2023. Notably, 101 studies have been published within the past five years (2019 – 2023).

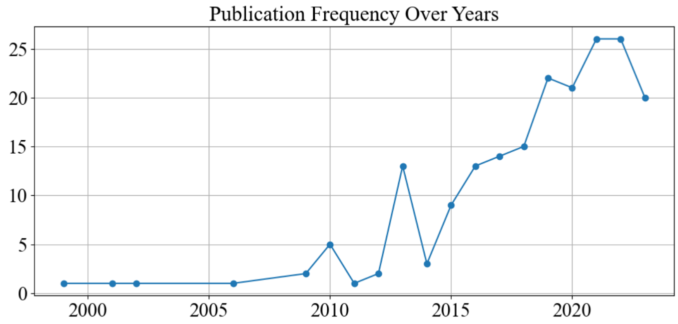


Figure 2.4: Publications over the years

Geographical origins The geographical origins of the studies are visually depicted in Figure 2.5. Notably, several publications resulted from collaborations among multiple organizations, totaling 202 instances. Western Europe accounts for 96 cases, followed by the United States with 40 instances, and East Asia with 37 cases. Together, these regions contribute to 77.2% of the overall publication instances.

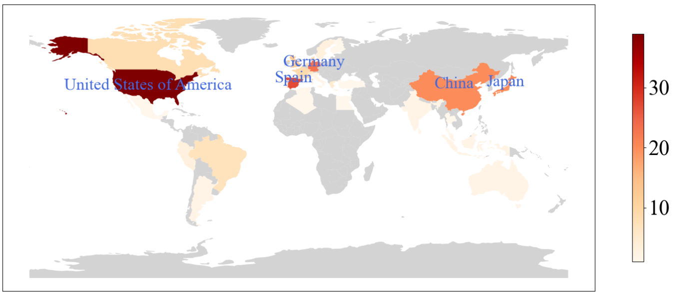


Figure 2.5: Global heatmap of the publications

Study Types Some publications involve both controlled and field settings, resulting in 197 distinct instances of settings. Of the studies examined, 89% (175 cases) were conducted in laboratories, whereas 11% (22 instances) occurred in field environments.

Participants Among the studies, 32% utilized only simulations without any real participants. Furthermore, 29% of the studies failed to mention participant demographics explicitly. Of the remaining, 20% involved fewer than 10 participants, 9% included between 11 and 30 participants, and only 10% involved more than 30 participants.

2.4. MAPPING SOCIAL CUES WITH INFORMATION

We analyzed human-robot information communication across all studies collected that yielded a quantitative overview shown in Figure 2.6, which includes numerical data on instances of social cues and information they communicate.

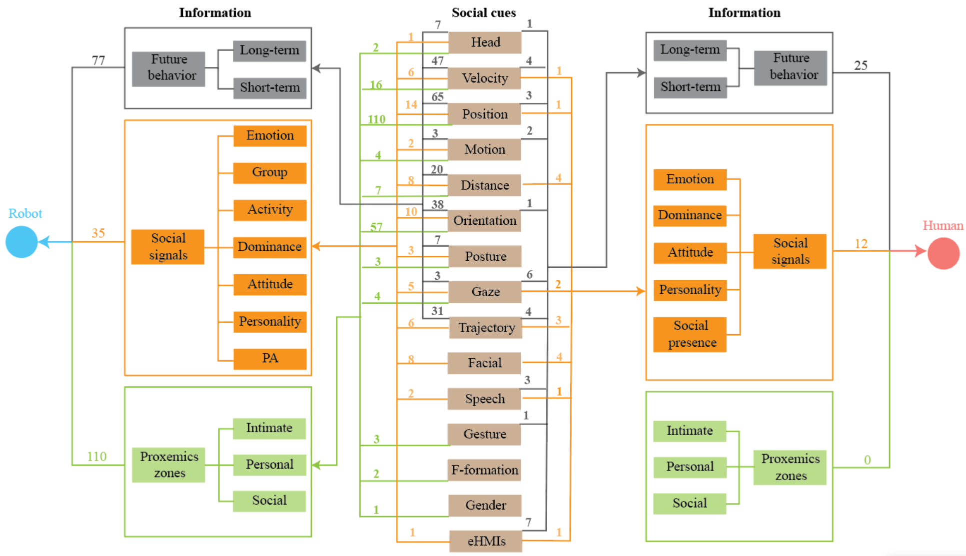


Figure 2.6: Quantitative data on human-robot information communication, where social cues serve as information channels.

On the left side, arrows from information to the robot indicate how the robot interprets human information, with numerical data showing total instances for each category: 77 for “Future Behavior," 35 for “Social Signals," and 110 for “Proxemic Zones." The arrows connecting social cues to this information indicate which human social cues the robot uses to understand human behavior, with numbers representing the total instances of these cues. Dominant cues for predicting human future behavior include “position", “velocity", “orientation", and “trajectory". For detecting social signals, robots mainly rely on “position", “orientation", “distance", and “facial expression". “Position" and “orientation" are also crucial for modeling proxemics zones.

On the right side, arrows from information to the human indicate the information that robots communicate to humans, with numerical data on total instances: 25 for “future behavior", 12 for “social signals", and 0 for “proxemic zones". Arrows connecting social cues to this information illustrate which robot social cues are used to communicate information to humans and include numerical data. Dominant cues used to indicate future behavior include “eHMIs", “gaze", “velocity", and "trajectory". For conveying social signals, “distance", “trajectory", and “facial expression" are most commonly used.

2.4.1. Predicting future trajectories

Most studies (n = 61) focused on robots predicting human paths, while a smaller number (n = 18) examined how humans predict robot behavior, known as robot motion legibility.

Robot predicting human trajectories

Predicting human trajectories is crucial for socially aware navigation and can be broadly divided into long-term goal and short-term trajectory predictions. Table 2.1 provides an overview of how robots incorporate social cues to predict human trajectories. Position and velocity emerge as the most significant cues utilized in 65 and 47 studies, highlighting their importance in predicting immediate and future locations. Orientation and trajectory are also notable, and they were used in 38 and 31 studies, respectively, indicating their relevance in assessing directional intentions. Less frequently considered cues include head position, posture, gaze, and motion patterns.

Long-term prediction, or goal prediction, involves forecasting the destinations of individuals within a given environment [58, 90–100]. These destinations, commonly referred to as "hot points" (HPs) [97, 99] or points of interest [98], include frequently visited locations such as doorways, desks, and coffee machines. These points can be annotated manually or identified through automated processes. In automated identification, a predictive map of these hot points is constructed offline, incorporating historical data on human trajectories, which includes both spatial and temporal aspects [90–92, 97, 98]. In addition, robots use nearby individuals' current positions and orientations, computing their most likely destinations. Various sophisticated algorithms support these predictions. Bayesian models [90–92, 95–97, 100], for instance, estimate the probabilities of individuals reaching specific destinations by integrating prior knowledge with real-time observational data. Other methods include D$*$ Lite [93], which provides a lightweight and efficient solution, Hidden Markov Models (HMMs) [98] that address the uncertainties inherent in human movement, mgIGP [99], and GoalFlow [58], which directs flow towards predicted goals. These predicted goals can enhance short-term predictions of human trajectories, such as in social force models where individuals are assumed to move towards a destination.

Short-term predictions involve predicting human trajectories within a short period, allowing robots to respond to immediate changes. These predictions involve various modeling techniques, including linear models, social force models, probabilistic forecasting models, and data-driven models. Linear models use linear assumptions about human motion, where velocity might be considered constant [101–109], or variable to accommodate more complex scenarios [110–112]. This approach relies on observable cues like position, orientation, and velocity. Social force models offer a more nuanced approach, modeling human behavior as responses to social forces exerted by the environment and other individuals. In these models, agents are influenced by attractive forces towards goals and repulsive forces to avoid obstacles, requiring the integration of social cues such as position, orientation, velocity, and distance to others [28, 44, 95, 96, 113–125]. Probabilistic forecasting models and data-driven models offer sophisticated techniques that cover both historical and real-time data. These models process trajectories along with current human positions and orientations [7, 25, 26, 29, 58, 60, 91, 92, 97, 126–148]. These models also consider the interpersonal distances [135, 149] and dynamically changing velocities [15, 150], postures [151], thus providing a comprehensive and predictive framework that adapts to the complexities of human behavior.

Table 2.1: Social cues for predicting human trajectories.

Predictions	Models	Head	Velocity	Position	Motion	Distance	Orientation	Gaze	Posture	Trajectory	References
				✓			✓			✓	[96]
			✓	✓		✓	✓			✓	[95]
	Bayesian							✓		✓	[97]
				✓						✓	[100]
Long-term				✓			✓			✓	[90, 92]
			✓	✓			✓			✓	[91]
	D∗ Lite			✓			✓			✓	[93]
	HMM			✓			✓			✓	[98]
	mgIGP			✓						✓	[99]
	GoalFlow		✓	✓		✓	✓				[58]
			✓	✓		✓	✓				[95, 113, 116, 117, 121]
			✓	✓		✓	✓		✓		[118]
		✓	✓	✓		✓	✓	✓			[114]
			✓	✓		✓					[115, 119]
	Social force		✓	✓		✓		✓			[120, 124]
		✓	✓	✓		✓					[44]
			✓	✓			✓				[96]
		✓	✓	✓	✓	✓					[122, 123]
			✓	✓			✓	✓			[28]
				✓			✓			✓	[25, 92, 134]
		✓	✓	✓							[29]
			✓	✓			✓				[91, 127, 129, 133]
			✓	✓		✓	✓				[58]
			✓	✓						✓	[7, 126, 138]
Short-term				✓		✓					[128]

Predictions	Models	Head	Velocity	Position	Motion	Distance	Orientation	Gaze	Posture	Trajectory	References
	Probabilistic									✓	[130, 131]
				✓			✓				[132]
						✓				✓	[135]
			✓	✓				✓			[136]
		✓	✓	✓						✓	[137]
				✓			✓			✓	[15]
	Data-driven									✓	[26, 135, 139–141]
					✓						[97]
			✓	✓						✓	[142]
						✓	✓				[143]
			✓	✓							[60, 125, 149, 150]
		✓		✓			✓		✓	✓	[144]
			✓	✓			✓				[146]
				✓			✓		✓		[151]
				✓			✓			✓	[145, 147, 148]
	Deterministic		✓	✓							[101, 103, 106, 108, 111]
			✓	✓			✓				[102, 105, 107, 110]
			✓					✓			[104]
		✓	✓	✓							[112]
			✓	✓							[109]
	Total	7	47	65	3	20	38	7	3	31	77

HUMAN PREDICTING ROBOT TRAJECTORIES

Table 2.2 provides an overview of how robots incorporate social cues to communicate intentions to humans. The most commonly used cues are eHMIs (external Human-Machine Interfaces) and gaze, appearing in 7 and 6 cases. Trajectory and velocity are also significant, each used in 4 cases. Other cues, such as head motion, orientation, gesture, and speech, are mentioned less frequently.

Socially aware navigation studies have employed various measures and terminologies, reflecting subtle differences in understanding human predictions. To address these differences, we categorized studies based on the terminologies used to describe human predictions of robot trajectories. Terms with different forms but the same root or meaning were grouped into the same category. For instance, "understanding," "understand," and "understandability" were all clustered under the category of "understandability," while "prediction" and "anticipation" were categorized under "predictability." The main terminologies clustered include "predictability," "legibility," "understandability," and "clarity" (See Table 2.2). Most studies have not explicitly mentioned or assessed human performance in predicting the long-term goals of robots, with the notable exception of the work by Hetherington et al. [17]. Regarding the social cues used, all studies utilized robot motion to communicate intentions, focusing on various motion-related features such as the robot's trajectory [31, 152, 153], position [30], velocity, and orientation [154–156]. Additionally, some studies incorporated robot gaze [32–35, 157], and various external interfaces to enhance communication of robot intentions. These interfaces included light signals [17, 33, 36, 155, 157, 158], wearable devices [159], touch devices [156], gestures [160], text or graphics [161], and sound [37].

Table 2.2: Social cues for robot trajectory predictions

Measures	Social Cues										References
	Head	Velocity	Position	Motion	Orientation	Gaze	Gesture	Trajectory	Speech	eHMIs	
Predictability			✓								[30]
				✓							[155]
									✓		[155]
						✓		✓			[34]
						✓					[32]
		✓									[37]
Legibility			✓			✓					[31]
		✓						✓			[152, 153]
										✓	[17]
	✓										[33]
							✓				[33]
						✓					[35]
Understandability			✓			✓					[157]
		✓									[154]
				✓							[156]
										✓	[156–159]
									✓	✓	[158]
									✓		[36]
					✓			✓			[160]
Clarity										✓	[161]
						✓					[32]
Total	1	4	3	2	1	6	1	4	3	7	25

2.4.2. PROCESSING SOCIAL SIGNALS

We identified 38 studies on social signal processing. Most studies (n = 31) focused on robots inferring human social signals, while a smaller number (n = 10) examined how humans infer robot social signals.

ROBOT INFERRING HUMAN SOCIAL SIGNALS

Table 2.3 presents a detailed examination of various social cues robots use to infer varied human social signals. These studies have predominantly focused on emotions, intentions, social relationships, and activities, focusing less on dominance, social attitude, and personality traits. The studies of emotion have been closely tied to facial expressions, though occasionally speech [55] or posture [162] have been used. While some research have focused solely on facial expressions for emotion detection [50, 51, 57], others have incorporated additional cues such as trajectories [42, 163–166]. The group relationship is often revealed through cues like position [105, 167, 168], orientation [162, 169–171], distance [107, 172], and gaze [94, 173]. Human activities such as walking or conversing have been linked to postures [174], or a mix of relative position, orientation [168, 175], and velocity [176]. Less explored social signals include dominance, modeled through motion features such as trajectory and velocity [52, 53], social attitudes such as politeness through speech [19] or motion [177], and personality traits detected from velocity and interpersonal distances [108]. Additionally, Perceived Appropriateness (PA) – a reflection of human feedback on robot behavior – can be inferred from head movements and human-robot distances [178].

Table 2.3: Social cues for inferring human social signals

Signals	Head	Velocity	Position	Motion	Distance	Gaze	Orientation	Posture	Trajectory	Facial	Speech	eHMIs	References
Emotion						✓		✓		✓			[51]
									✓	✓			[42, 163–165]
											✓		[55]
								✓					[166]
										✓			[50, 57]
								✓					[162]
Group	✓	✓	✓		✓								[107]
												✓	[179]
		✓	✓		✓					✓			[172]
			✓			✓	✓						[94]
					✓								[167]
			✓		✓		✓						[170]
					✓								[105, 168]
			✓				✓						[28]
			✓	✓									[26]
					✓		✓						[169, 171]
			✓			✓	✓						[173]
			✓				✓						[162]
		✓	✓										[180]
Activity		✓	✓				✓						[176]
			✓				✓						[168, 175]
								✓					[174]
Dominance				✓									[53]
		✓							✓				[52]
Attitude									✓				[177]
											✓		[19]
Personality			✓			✓							[108]
PA		✓				✓							[178]
Total	1	6	14	2	8	5	10	3	6	8	2	1	35

Human inferring robot social signals

Table 2.3 presents a detailed examination of various social cues robots use to communicate social signals. Different signals require distinct cues: emotions are mainly conveyed through facial expressions [54–57]; social presence and attitude through trajectories and distance to humans [29, 181–183]; personality traits through velocity and distance [184]; and dominance through trajectory [53].

Table 2.4: Social cues for communicating robot social signals

Signals	Velocity	Position	Distance	Gaze	Trajectory	Facial	Speech	eHMIs	References
Emotion						✓	✓		[55, 56]
			✓						[20]
						✓			[54, 57]
Social presence			✓	✓					[182]
					✓				[20]
				✓					[185]
			✓					✓	[181]
Attitude		✓			✓				[29]
Personality	✓		✓						[184]
Dominance					✓				[53]
Total	1	1	4	2	3	4	1	1	12

2.4.3. Estimating proxemics space

Accurately estimating and modeling personal space is fundamental to facilitating natural and comfortable human-robot interactions. Among the collected studies, 110 works have investigated proxemics space, all from the human perspective. Each space category is critical in defining how robots should navigate around humans, maintaining distances and behaviors perceived as socially acceptable and comfortable.

Table 2.3 presents a detailed examination of various social cues robots use to model human proxemic spaces 2.5. Most of these studies—84 in total—focused on personal zones. In comparison, 17 studies modelled social zones and nine intimate zones. The primary social cues for modelling proxemic spaces are a person's position and orientation. Additional cues, such as velocity and distance, also contribute to the modeling of proxemic space.

Intimate zones were traditionally modeled as circular regions [44–46, 168, 186], focusing on an individual's position and assuming uniform distance in all directions. Researchers have proposed elliptical models to reflect the directional nature of human interactions better. These models, using asymmetric Gaussian functions, consider both the position and orientation of a person [14, 122, 187]. The elliptical shape extends more space in front and less behind, reflecting our forward-focused behavior and aligning more closely with real-world interactions and cultural norms.

Personal zones, pivotal to social interactions, have been visualized in geometric shapes such as circular, elliptical, or asymmetrical. Hall's model defines personal space as a circular region extending from 0.45m to 1.22m around a person, a foundational concept widely accepted in numerous studies [38]. This model emphasizes the significance of an individual's position in space [15, 28, 44, 46, 95, 101, 115, 117, 124, 130, 186, 188–199]. Social signals, such as personality [108] and emotions [50, 163, 164, 166, 200], also influence personal space. Traditional circular-shaped models may not accurately reflect the nuanced dynamics of human social behavior. Other studies have modelled personal space as elliptical shapes [114, 118, 122, 175, 201, 202]. These models, which are either (semi-) elliptical, consider both the position and the orientation of an individual [103, 203, 204]. Beyond elliptical models, researchers have explored the asymmetric nature of personal space. These models primarily rely on an individual's position and orientation, effectively differentiating the space in front, which is typically more extensive, from that behind [14, 39–41, 45, 54, 93, 100, 168–170, 176, 179, 205–215], and can be influence by emotional states [151]. Furthermore, these Gaussian-based models have incorporated social cues of velocity and hand positions [16, 32, 92, 106, 132, 133, 158, 171, 180, 187, 191, 216–220], human demographics (height, size), and density of the surrounding area [221–225].

Social zones were also clustered into three shapes: circular, elliptical, or asymmetrical. The concept of a circular social zone, deeply rooted in the seminal work of Hall [38], typically encompasses radii ranging from 1.2m to 3.6m, which are considered optimal for social interactions [96, 115, 186, 196, 198]. An individual's position fundamentally influences such zones and is further affected by personal attributes like personality traits [108] and emotion [43, 164]. Elliptical zones represent a more sophisticated modeling approach, moving beyond circular designs. Dominantly shaped by Gaussian functions, these zones incorporate an individual's position and orientation, velocity [226], human-robot distance [227], as well as head and hand positions [191]. This shape effectively captures the directional nature of human attention and interaction, providing a more accurate framework for understanding and navigating human spaces. Others have also proposed asymmetrical shapes, which are especially relevant in environments where multiple individuals interact, requiring an understanding of each group member's position [16, 191, 212, 228, 229]. By incorporating socio-spatio-temporal characteristics, such as human velocity, these models offer a dynamic and socially aware mapping of social zones.

Table 2.5: Social cues for modelling humans' proxemics space

Proxemics	Social Cues											References
	Head	Velocity	Position	Motion	Distance	Orientation	Gaze	Gesture	F-Formation	Posture	Gender	
Intimate			✓									[44, 46, 106, 168]
			✓			✓						[14, 45, 122, 186, 187]
			✓									[15, 28, 44, 46, 50, 95, 100, 101, 108, 115, 117, 124, 130, 163, 164, 166, 168, 171, 188–192, 194, 195, 197–200, 211, 213, 216]

Proxemics	Social Cues											References
	Head	Velocity	Position	Motion	Distance	Orientation	Gaze	Gesture	F-Formation	Posture	Gender	
Personal			✓			✓						[14, 41, 45, 54, 93, 103, 114, 118, 122, 132, 151, 158, 169, 170, 175, 179, 186, 187, 193, 196, 201, 203–205, 208–210, 214, 217, 219]
			✓		✓					✓		[206, 215]
		✓	✓		✓	✓	✓					[220]
		✓	✓		✓	✓					✓	[39]
			✓						✓			[176]
		✓	✓		✓	✓				✓		[92]
	✓	✓	✓									[106]
		✓	✓									[16, 180, 221, 225]
			✓			✓			✓			[40]
			✓	✓		✓	✓	✓				[191]
			✓			✓	✓					[212]
		✓	✓			✓						[202, 207]
		✓	✓									[224]
		✓	✓			✓						[133, 218, 222, 223]
			✓					✓				[32]
Social			✓									[43, 95, 96, 108, 115, 164, 198]
			✓			✓						[196, 212, 213, 228, 229]
					✓	✓						[186]
		✓	✓		✓	✓						[227]
			✓			✓	✓					[28]
	✓		✓	✓		✓		✓				[230]
			✓	✓		✓						[226]
Total	2	16	110	4	7	57	4	3	2	3	1	110

2.5. DISCUSSIONS AND SUGGESTIONS

We conducted a scoping review on human-robot information communication in a navigation context. We systematically searched studies in robot social navigation and analyzed the information communicated and social cues used for communication in the 176 studies collected. This resulted in a comprehensive list of social cues and three distinct clusters of information: future states, social signals, and proxemic zones. We further provided a quantitative overview of social cues and information they communicate to inform human-robot interactions.

As shown in Figure 2.4, an increasing number of publications over the last two decades indicates a significant interest in socially aware navigation. This field is predominantly influenced by research from Western Europe and the United States, with minimal contributions from regions such as Africa, Central Asia, and South America, as shown in Figure 2.5. However, studies revealed strong cultural variations in social norms such as personal space requirements[151, 165, 203] and human walking habits [33], which fundamentally shape the expectation of robot navigation behavior. Therefore, we propose **Future research should prioritize studying cultural variations in social norms, especially in underrepresented regions like Africa, Central Asia, and South America.**

In the collected studies, 88% were controlled, while only 12% were field studies. Some studies chose controlled lab settings to focus on specific variables. For example, a straight navigation path was used to systematically analyze how variations in robot velocity affect human perception of robot motion [154]. These controlled settings also enabled the assessment of how trajectory features and socio-contextual constraints improve the legibility and social acceptance of robots by minimizing external variables [160]. On the other hand, some studies were conducted in field settings, acknowledging that effective social navigation in urban environments necessitates real-world testing [46, 96, 231]. Field tests address the biases inherent in controlled studies and support real-life applications [33, 127] but are currently much less studied. Consequently, we propose **Suggestion 2: More field studies should be conducted to enhance the ecological validity of the research outcomes.**

Among the collected studies, 32% involved simulations, while 29% of real-life experiments lacked detailed demographic information about the participants. Moreover, among the studies that did provide such details, the majority included fewer than 11 participants. Although simulations are helpful for initial hypothesis testing and controlled experimentation, their generalizability is often limited. Simulated human participants follow basic rules in navigation, which may not accurately reflect the complex human behaviors and interactions observed in real-world settings. Many studies have employed simulations and real-life experiments to validate their findings [60, 179, 225]. However, the absence of detailed demographic information in a substantial portion of real-life studies further compromises their generalizability. We propose **Suggestion 3: Future studies should enhance the robustness of their findings by increasing both the diversity and number of participants**.

Our survey reveals a predominant focus on how robots interpret human information, accounting for 83% of the research, compared to how robots communicate information to humans, which constitutes only 17%. Advancements in algorithms and models have enhanced robots' capability to understand human behavior: predicting human movement that allows robots to avoid conflicts and collisions [107, 124]; modelling and respecting personal space ensure that robots respect human comfort and safety [39, 218]; while detecting human social signals, such as activity, emotion, and dominance levels, enable robots to better adapt to individual preferences [55, 184]. How robots communicate their states and intentions to humans is less explored. Robots that exhibit greater legibility in their motion can significantly increase their social acceptance [17]. Furthermore, robots conveying social signals, such as emotions, are facilitating more realistic interactions [55]. Therefore, we propose **Suggestion 4: We should conduct more studies on how robots communicate information to humans, such as emotions, intentions, or politeness.**

Robots could improve the predictability or legibility of their motions through various cues, including motion features and external interfaces such as lights, sounds, and text (Table 2.2). Notably, most research has focused on the short-term predictability of robot motions, with limited emphasis on long-term intentions, except using lights to signal a robot's long-term goals [17]. Gopalakrishnan et al. [18] argue that short-term cues are generally sufficient for human interaction because humans quickly adapt to immediate environmental changes. However, Hetherington et al. [17] provide evidence that understanding a robot's long-term goals can significantly improve human prediction of robot trajectories, thus enhancing overall interaction performance. This highlights a research gap that merits further exploration into how long-term and short-term cues might be combined to improve the legibility of robot motion. Therefore, we propose **Suggestion 5: Future studies should consider long-term cues in robot trajectory communication to enhance human understanding.**

Despite studies on human-robot communications (Table 2.3 and Table 2.4), there remains a notable deficiency in specific types of social signals, especially social attitudes, dominance levels, and the perceived appropriateness (PA) of robot navigation behavior. As proposed by Jered et al. [178], PA highlights a crucial component of social navigation—the feedback on the appropriateness of a robot's behavior. This type of feedback is essential for robots to adapt their behavior to individual preferences. Therefore, we propose **Suggestion 6: Future research should expand on studies of diverse social signals such as dominant levels and social attitudes, especially human feedback.**

Current studies on proxemics all focus on how robots model human proxemics spaces instead of robot proxemics (Table 2.5 and Figure 2.6). According to the Computers-are-social-actors paradigm [4, 5], humans interact with robots and other artificial agents using social behaviors similar to those they use with other humans, and it is likely that humans also perceive proxemics zones in robots. Key questions include whether humans recognize a proxemics space for robots and how they perceive its size relative to their own. Addressing these questions is crucial for designing robots capable of effectively managing spatial relationships with humans. Furthermore, most existing studies utilize Edward T. Hall's definitions of human proxemics [38], which may not fully apply to interactions involving robots, given the physical and behavioral differences between humans and robots. Investigating how various robot attributes influence human expectations and comfort levels regarding proxemics is thus essential. Therefore, we propose **Suggestion 7: More studies should be conducted to understand human perceptions of robot proxemics spaces.**

Current literature on socially aware navigation has begun to explore how information influences each other, yet a comprehensive understanding of such influence remains limited. While some studies have identified the effects of emotions and personality traits on individuals' personal space (proxemics) [19, 163, 165], these dynamics are not thoroughly investigated. A more detailed study of these interactions could provide critical insights, enabling robots to understand better and respond to human behaviors. If robots could accurately detect human emotions and personality traits, they might more accurately predict human trajectories, facilitating smoother interactions. Consequently, we propose **Suggestion 8: Future studies**

should investigate how different types of information influence each other.

2.6. Conclusion

This review has systematically examined human-robot communication in socially aware navigation studies, focusing on the types of information communicated and the social cues used. The analysis of 176 studies has yielded a comprehensive taxonomy of social cues and information types, providing a critical framework for advancing socially aware navigation research. Our thematic analysis has resulted in developing a model that categorizes three distinct types of information: future behaviors, social signals, and proxemics spaces. We further conducted a detailed analysis of every kind of information. We provided quantitative data on the prevalence of these communication types in the surveyed studies, revealing a dominant focus on how robots interpret human behavior instead of human perceptions of robots. Based on these findings, we recommend that future research explore human perceptions of robot behavior—particularly regarding feedback information such as perceived appropriateness (PA)—to enhance cooperative navigation. Moreover, our review has provided a quantitative overview of the use of social cues in communicating different information. These insights facilitate the design of more informed and effective robot behaviors, thereby contributing to integrating robots into human societies.

3

Human Yielding Behavior and Perceptions

The increasing presence of mobile robots in public spaces necessitates seamless navigation among pedestrians. However, robots often fail to navigate around humans due to a lack of understanding of the dynamic back-and-forth involved. We thus aim to answer the research question: What are the key factors contributing to human yielding to conflicts in public spaces, and how do they influence human PA of robot navigation behaviours (RQ2)? Based on a pre-study, we developed observation forms to track various factors and human responses without imposing on people's privacy. We then conducted a field study (427 encounters observed) with a teleoperated robot deliberately ignoring humans. Our observations and follow-up interviews highlight various vital factors that jointly relate to human-yielding behavior, such as the setting (wideness of the passageway), robot behavior, human states (being stationary, teasing, cycling, carrying items, playing with cellphones, assumed age), and perceived appropriateness (PA). We thus pinpoint three key design considerations for social navigation: empathic awareness, PA responsiveness, and readable expressiveness. Together with the identified factors and the novelties in our method to support extracting these kinds of insights, this work contributes to the design of and future research into making robot social navigation more streetwise.

3.1. Introduction

Robots are increasingly navigating social spaces for various applications, including delivering goods [232, 233], inspecting or cleaning public areas [234], guiding visitors [235, 236], and providing information [237, 238].

However, the challenges of navigating dynamic and complex real-life situations are still not fully understood (e.g., [79]). The pilots conducted thus far (such as

This chapter has been submitted as: Yunzhong, Z., Vroon, J., Zoltan, Z., Gerd, K., 2024, Streetwise Social Navigation: Field Observations of Yielding in Robot-Pedestrian Interactions.

Starbot's food delivery robots) highlight the real-world issues that can occur when autonomous robots navigate among pedestrians, including a wheelchair user being blocked by robots [239] and robots being assaulted and bullied [240–242]. The increase in such incidents has led some cities to enact partial bans on autonomous delivery robots (e.g., [243]).

Yet, despite a plethora of research on robot navigation, there is but limited research on how it works or breaks down in real-world incidental encounters [244–249]. Existing navigation strategies, including the social force model and its variations [76–78], can still exhibit inappropriate behaviors. These behaviors may be tolerable at times but become problematic when they constantly force humans to yield to robots, contradicting socially aware navigation principles of least disturbance. While yielding behaviors and their influencing factors have been explored in other contexts, a.o. between pedestrians and drivers [250–252], our understanding of yielding between robots and pedestrians in public spaces such as sidewalks is still lacking. To quote the recent conclusions of Gao and Huang [11], there is a need for "more [...] field studies to productively advance socially-aware robot navigation and develop useful, functional mobile robots" (p.13) and a need to "enrich pedestrian models" and "measure and report individual characteristics" (p.14-15).

We thus investigate the research question: What factors influence the yielding behavior of pedestrians when encountering a robot in a public space? As they are what triggers and necessitates yielding behaviors, we focus on **conflicts**, following the definition of Mirsky et al.: "A conflict between a robot and [...] pedestrians is a situation in which if there is no change of direction or speed by at least one of the parties, they will collide" [79, p.2]. Rather than exploring the appropriateness of specific social navigation strategies, we aim to understand how potentially inappropriate situations develop in the wild. We focus specifically on pedestrians and factors directly relating to their behavior, leaving factors specific to the robot out of scope. To answer the research question, we conducted a field study with a Clearpath Husky robot in an outdoor public space in a bustling area with both narrow sidewalks and a more open pedestrian zone. We used context confrontation [253], a form of field observation developed to understand the natural reactions of pedestrians when a robot is seemingly navigating (seemingly) without responding to humans. This means we did deliberately not imbue the robot with any social behavior; the navigation ignored all humans. A robot operator fully controls the robot (speed, orientation, direction) to allow encounters to emerge naturally but safely. We developed detailed observation forms in a pre-study to enable rich and systematic naturalistic observations while avoiding infringing on people's privacy. We further enriched this with a sampling of interviews, allowing us to explore human perception of the robot's behavior and various social cues more deeply.

Our findings indicate that most often, pedestrians yield for a robot and avoid potential conflicts a long time before collisions would have occurred. On the other hand, there were also cases in which pedestrians did not yield or did so at the last minute. As the main outcome, we identify various factors that correlate with this yielding or non-yielding behavior (Table 3.1) and derive design considerations from them. Our findings have several implications for robotics research. First, they highlight a range of situations where robot social navigation should be designed more carefully to be safe and acceptable, i.e., “streetwise”. Second, they demonstrate that field observations such as the context confrontation we conducted—utilizing observation forms and a robot deliberately not employing any social strategies—play an essential role in systematically investigating the requirements for social navigation strategies in specific situations.

3.2. RELATED WORK

Social navigation among pedestrians is complicated when considering the number of entities involved. Yet, we mostly manage to avoid each other without collisions [254, 255]. Much of our understanding of this social positioning behavior is based on early sociological work investigating our personal space, as described by Hall's proxemics [256, 257], and the different formations that groups of humans tend to form (see, e.g., Kendon's work on F-formations [258]).

Still, social navigation encompasses many different activities: we can walk side-by-side (or in other formations) (e.g., [78]), we can approach one another (e.g., [237]), we can position ourselves relative to each other concerning each other's personal preferences, and we can pass each other by while mostly avoiding collisions (e.g., [247]). Within the context of this work, and given our subsequent focus on incidental encounters [249] and conflicts [79], we will focus on this latter case of passing by, the conflicts that play a role in it (e.g., [259]) and the yielding behavior of pedestrians.

3.2.1. SOCIALLY-AWARE ROBOT NAVIGATION AND CONFLICTS

From a robotics point of view, social navigation entails numerous technical challenges that have all been worked on extensively. Many of these works start from applicable sociological concepts of social positioning in human-human interaction and try to formalize them for application in robotics (as reviewed by, e.g., [260, 261] and reflected on by, e.g., [11]). As a consequence of this, social navigation is often phrased as trying to navigate such that the robot never gets closer to people than certain set distances—with those distances being derived from sociological concepts like Hall's personal space zones and F-formations (see, e.g., [262]). Based on this definition, a series of applicable technical challenges will be tackled, from computer vision to predicting pedestrian behavior to path planning (e.g., the review by [259]).

While pragmatic and relatively effective, such approaches pose challenges and trigger a range of conflicts (e.g., the recent review by [79]). For example, how should we balance avoiding such intrusions against deviating from the shortest path [263], and how should we weigh multiple intrusions against each other [264]? Another complication with this approach is that people often respond to the behavior of the robot—which has led developers to look into the legibility of their navigation behavior [265] and into ways in which these response behaviors can be used to further the navigation [263, 266] or change people's proxemics preferences [266]. Recent work has also looked into using relevant social cues (e.g., [64, 79]) to detect such conflicts as they arise [267], which is fundamental to enabling robots to respond appropriately.

Such challenges also affect robots deployed in the wild, hindering their acceptance in our social spaces. From numerous examples of (mobile) robots being harassed and attempts to mediate this [240–242], to mobile robots bothering (vulnerable) road users (e.g., [239], the robot crossing a crime scene, and the partial ban of mobile robots in San Francisco [243]).

3.2.2. Robots in the wild; field observations for social navigation

Field observations play a significant role in human-robot interaction, offering essential insights into robot design [268]. These studies examine specific contexts and human reactions, such as exploring people's reactions to a robot trash can [269], conducting trials with a robot providing information in a shopping mall [238], and investigating the social positioning behavior of people around a robot in an indoor setting [270]. However, field studies on social navigation have been limited, with most research concentrating on evaluating specific solutions (see, e.g., [246, 248, 271]). However, how much of their observations can be attributed to designed robot behavior is unclear. More recently, there have been three independent works, all of which have been conducted using field observations. Van Mierlo conducted exploratory systematic observations into the reactions triggered by a delivery robot [272]. Still, he focused primarily on mapping out the types of responses instead of the underlying factors. Both Ambrams [273] and Vroon [253] reported a few fascinating cases of people responding to such robots, but these are based on preliminary and mostly informal observations. As such, we still lack a more systematic exploration of robot social navigation in context, specifically in the relevant factors that should be considered.

3.3. Study

We aimed to investigate the factors influencing pedestrian yielding behavior to a robot in an outdoor public space. To this end, we needed a set-up with naturally occurring conflicts between pedestrians and a mobile robot. We conducted field observations with unwitting pedestrians—using observation forms rather than video recordings to avoid privacy issues.

In this chapter, we will give an overview of the location (Chapter 3.3.1) and the robot behaviors (Chapter 3.3.2) we used to conduct these observations. We will then discuss how we developed our observation form with a pre-study (Chapter 3.3.3) and how we set up a small convenience sampling of interviews to collect information on how pedestrians perceived the behaviors of the robot (Section 3.3.4). All these components come together in the procedure for our systematic observations, which had to balance various ethical concerns (Chapter 3.3.5).

3.3.1. CONTEXT

We conducted the study at the *Marineterrein Amsterdam Living Lab (MALL)* (see Figure 3.1). This is a publicly available area, close to the city center of *Amsterdam*. The location encompasses narrow sidewalks and wider streets primarily used by pedestrians (cars are discouraged from entering). We further conducted observations on a sidewalk next to a cycling path.

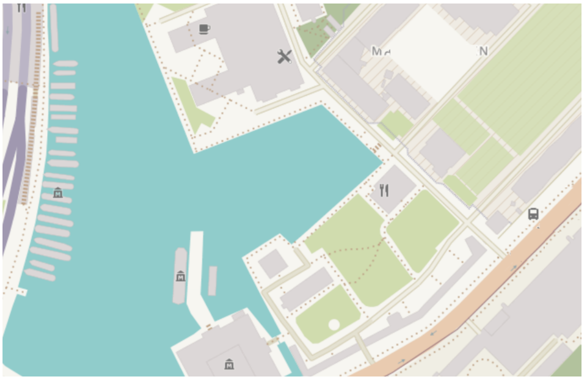


Figure 3.1: Map of the *Marineterrein*, where we conducted our field observations. © OpenStreetMap contributors

The location is used as a through-path for pedestrians and cyclists, as a calm area for walking dogs, as a destination due to the various offices, restaurants, cafes, inn, and outdoor gym, and as a recreational zone, with its small park and public waterfront pool in the former harbor. This meant that our observations included everything from people calmly strolling or walking to people running or cycling more rapidly toward destinations. The location is still privately governed as a former navy base, giving it a unique breeding ground for studies into the future of the city.

Pedestrian traffic during the study varied from quiet mornings to more busy lunch times and extremely busy, sunny afternoons. Though we conducted the first round of observations during the COVID-19 lockdown measures (summer 2021), we will focus on a second, extended round of observations under much more regular circumstances (summer 2022). We mainly observed on sunny days (and during some light drizzle) but avoided rainy days to protect the robot.

3.3.2. MATERIALS

We used Clearpath Robotic's Husky as our robot platform for this study. It is comparable in size to many of the sidewalk robots/vehicles currently under development, albeit somewhat less tall (dimensions [bwh]: 990 x 670 x 390 mm, 50kg). For safety and convenience, we conducted our study with complete manual robot control by a somewhat hidden human operator (Wizard-of-Oz set-up).

Protocol for the robot's behavior As discussed, we aimed to allow conflicts to emerge without deliberately causing them. To get clear observations of the yielding behavior of pedestrians in response to these conflicts and to avoid biasing our findings by introducing new factors, we explicitly did not give the robot itself any yielding behaviors—unless doing so was necessary for safety reasons.

This resulted in the following protocol for controlling the robot. (1) Drive on one side of the road/sidewalk at a speed of 1.0 m/s (comparable to a slow walking speed), (2) do not deliberately aim toward humans, (3) drive around all non-human obstacles, such as cars and parked bicycles. (4) Ignore all humans; do not drive around them or slow down. (5) When at risk of colliding with a human (i.e., when a pedestrian is within 50cm in front of the robot), make an '*emergency stop*' for safety reasons. Continue driving if/when the path is cleared, but stick to the initially planned path. (6) If common sense or safety dictates it, abandon protocol and note this down explicitly.

For safety and due to the range of the Bluetooth controller, the robot operator had to stay within approximately 10 meters of the robot constantly. This meant the operator had to walk near the robot visibly; to not be directly associated with the robot, we employed various techniques, including not constantly looking at the robot, not walking at the same speed as the robot, walking across the street from the robot, and holding the controller at waist-height.

3.3.3. Pre-study to inform our observation forms

To inform the set-up, we first conducted informal observations of interactions between human pedestrians to understand the factors that might play a role. To do this, we sat in various public spaces (in Delft and our campus). We paid specific attention to conflicts between pedestrians, i.e., situations where a collision would occur if not one of the involved parties would yield. After that, we used the same procedure to conduct the first informal observations with a small prototype 'robot' made out of cardboard and a remote-controlled car, using the protocol for behavior outlined above. Besides this, the pre-study observations followed the same protocol as the study-as-a-whole.

We made notes throughout and used a grounded theory approach to develop themes based on the patterns we observed iteratively. Our observations were primarily conducted by one observer, who then discussed the patterns to find common ground. These steps follow the method of context confrontation [253] These initial steps are discussed in more detail in our earlier publication on the method [253].

To enable us to capture observed interactions between the robot and pedestrians quickly and more systematically, we then developed observation forms (Figure 3.2) using the observed patterns and some basic demographics (assumed age, assumed gender) as a guideline. We iterated the layout of the forms to create a more efficient and natural flow. In the remainder of this section, we will briefly discuss the patterns we observed during the pre-study and how they informed the concepts and fields we used in our observations and our observation forms.

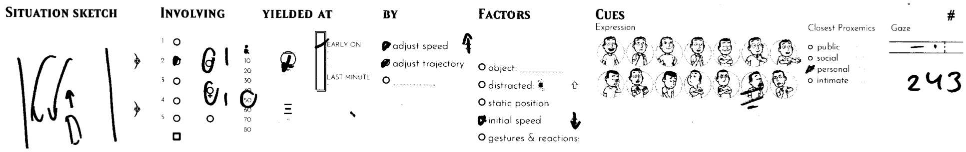
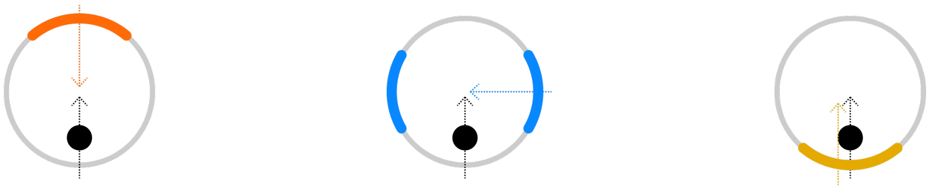


Figure 3.2: Filled in the example of the observation form we developed for and used in our observations. Following the elements from left to right, the situation sketch shows a heads-on conflict in a mostly straight wide stretch of road involving the robot and a group of two pedestrians, assumed to be a male and female of about 50 years old, who yielded for the robot, at early on, by speeding up and adjusting their trajectory. The only applicable factor was that their initial speed was somewhat slow. In terms of cues, they very strongly showed boredom (annotations based on the PrEmo tool [274]), got no closer to the robot than the personal space zone [256], and looked at it but a few short times before passing it and not after passing.

Conflicts During our pre-study, we observed three distinct categories of conflicts based on the angle between the projected headings (see Figure 3.3): heads-on, crossing, or from behind. All conflicts we observed belonged to one of these categories, which may be due to our public spaces' design.



(a) Heads on (b) Crossing (c) From behind

Figure 3.3: Conflicts occur when the projected headings of two road users intersect. We distinguished conflict types by the relative angle between these headings: *heads on* for opposed directions (including heading towards stationary pedestrians), *crossing* for orthogonal directions, and *from behind* for aligned directions. In rare cases, non-straight headings also intersected, e.g., when both the robot and a pedestrian would turn to go in the same direction at an intersection.

Yielding behavior When looking at yielding behaviors during our pre-study, three aspects meaningfully described them: whether or not the pedestrian(s) yielded at all, when they yielded, if they did, and how they yielded.

When pedestrian(s) did not yield at all, since the robot did not yield either, this always resulted in an 'emergency stop', making this easy to observe. Our initial observations distinguished three types of yielding behaviors: changing speed, changing trajectory, or something else (e.g., stop and wait).

The timing of the yielding behavior was diverse and (thus) hard to capture precisely. We saw timings ranging from people who yielded for each other and our prototype robot while still over 15 meters (and several seconds) away to people not yielding until the last second within decimeters of collision). In an early exploration of the study design, we found big annotation differences between observers. We thus explicitly aligned our observers by taking the aforementioned two examples as the extremes of 'very early on' and 'very last minute', respectively, and defining the scale as continuous.

Factors In our initial observations, we built a first intuition about different factors that might play a role in yielding behavior, i.e., activities (e.g., walking dogs, cycling, carrying items), distraction (e.g., cell phones), whether or not the pedestrian(s) was moving, initial speed. While not excluding other factors—we left ample space for comments—we added these factors to our form for easier and more consistent annotation.

Cues and perceived appropriateness We further observed various social cues pedestrians use in their interactions, most notably facial expressions, proxemics, and gaze. We thought these aligned with their perception of the robot, mainly whether they felt it should adapt its behavior.

We also included the first constrained coding of these cues to explore this relation further. We annotated expressions using the PrEmo tool, as it provides an accessible set of 14 emotions related to product experience [274]. For proxemics, as reliably capturing distance between the robot and pedestrians over time would be unfeasible with an observation form, we focused on just recording the minimum of that distance for each encounter, i.e., the closest proxemics distance to the robot [256]. Similarly, to track gaze feasibly, we used rough coding of gaze frequency and duration based on a timeline crudely indicating when pedestrians looked at the robot.

A sampling of follow-up interviews was used to collect the impression of pedestrians (see below, Section 3.3.4).

Demographics To get a rough idea of how representative our sample was, we recorded basic demographics (assumed age category, assumed gender) on all people observed in conflicts with the robot. As these demographics are based only on a rough and superficial assessment by the observer, we explicitly tested for the (inter-rater) reliability of this coding to confirm they were sufficient for our purpose.

3.3.4. Interviews

As our observations alone did not give any information on how the observed pedestrians subjectively experienced these interactions, we enriched our observations with a subsampling of interviews.

After some encounters during the study ($\approx 7.5\%$), we immediately approached the involved pedestrian(s) and conducted a brief semi-structured interview, asking them to reflect on the encounter. If multiple people were involved in the encounter, we interviewed them as a group. To cover a range of perspectives, we explicitly aimed to widen the range in age, gender, group size, and yield behaviors; otherwise, our sampling would have been convenience-based. Some pedestrians declined our request for an interview, particularly during one exceptionally warm day.

Our interview contained three questions, asked in Dutch or English, depending on the preference of the interviewee(s). An open-ended question was used to probe our interviewees' experience in a non-suggestive way: "You just encountered a robot; how did you experience that?". Then, we asked participants to pick the icon that best represented their emotional state from the PrEmo tool that we also used in our observations; this provided a baseline to check our observations against and additionally served to prime our participants for a more effective reflection. The last question was more directive, asking for suggestions for improvement to pinpoint if/where the robot's behaviors were seen as inappropriate: "Should the robot adapt his behavior to feel more appropriate to you? If so, how?"

3.3.5. Systematic observations

Our procedure for the systematic observations follows relatively straightforwardly from what was discussed above. In the summer of 2022, two observers visited the location to conduct our observations. One used the robot and protocol described to navigate back and forth around the terrain for a few hours, taking breaks as required (a.o. to replace the battery). The other observer would not down any conflicts using the observation forms. The two observers switched roles regularly. Other salient occurrences were also reported. Both observers annotated one cluster of observations to check the procedure and investigate inter-rater reliability. In the first session, the two observers were explicitly trained in using the observation forms, focusing on aligning them with subjective factors such as yield timing.

Ethical considerations As argued above, it was fundamental to this work to make these observations with unwitting pedestrians in the will, -yet this raised several ethical concerns. While making video recordings would have been convenient, it would have invaded the privacy of those pedestrians. Our approach with observation forms provided a suitable alternative to making systematic observations.

Furthermore, deliberately allowing conflicts with unsuspecting bystanders to emerge could pose a risk. While most of these 'conflicts' were effortlessly and harmlessly resolved by the pedestrian yielding, we still endeavored to reduce further the risk by manual control and constant supervision of the robot, including safety precautions in the protocol and using a relatively small and light robot. In addition, the location where we conducted our study Marineterrein, is marked as a field lab at all entrances.

With these considerations, this study was evaluated and approved by the ethical committee of the University of Delft.

3.4. Findings

Using the developed observation form, we conducted well over 30 hours of systematic observation, spread out over two weeks, during which we observed 427 distinct conflicts with the robot.

After a discussion of inter-rater reliability in our observations, we will here discuss the critical patterns of the conflicts we observed (Section 3.4.1), patterns in the relation between social cues and perception of the robot (Section 3.4.2), as well as more incidental other observations (Section 3.4.3). An overview of the main factors we report on can be found in Table 3.1. For these patterns, we looked at all aspects of yield behavior as defined above; yield timing, particularly, had a lot of variance between encounters. For that aspect, we were able to look deeper into the interplay of related factors. As the markings for yield timing were not precise enough to represent a continuous number (for example, the slanted mark in Figure 3.2), following discussion with the observers, we binned the annotations on a 7-point scale from ‘very early on’ (1) to ‘very last minute’ (7).

Our work aimed to identify a wide range of potentially meaningful factors, not to formally test their impact. Therefore, we will present and describe distributions instead of using evaluative statistics; as the patterns were identified post-hoc based on our data, we want to avoid explicitly suggesting evidence for assumed causality. To check the reliability of our observations, we conducted one session where both observers coded the same 20 conflicts. We only did this for 20 conflicts, as it required notably more focus from the observer driving the robot; we also swapped roles halfway for these 20 conflicts to reduce any potential effect on our reliability analysis.

There was no confusion between conflict types (agreement 100%). The observers fully agreed on whether or not the pedestrian(s) yielded (100%). The yield timing was acceptable to good at the last minute or early on (Pearson’s r 0.75). How pedestrians yielded was also reasonably well-aligned, with the observers agreeing in most cases on both speed (14/20, 70%) and trajectory changes (19/20, 95%).

Assumed gender was agreed on in most cases (29/31), and agreement on assumed age was good (Pearson’s r 0.92). Coding of the social cues from observations yielded poor agreement for expressions (9/20) and reasonable agreement on the proxemics zone (Spearman’s ρ 0.58). Gaze duration and frequency did not have enough variation in this sample to justify testing for agreement. Beyond the cues, these findings suggest a sufficiently high inter-rater agreement, justifying further analysis.

3.4.1. Conflicts

We observed 427 conflicts, at roughly a conflict every 4 minutes. Barely any of these ‘conflicts’ were associated with a strong outward emotional reaction (34 very strong expressions, albeit not necessarily in response to the robot). The vast majority of conflicts we observed were heads-on (69%). From-behind conflicts, where pedestrians and cyclists overtook the robots, were the second most common (17%), and crossings were the rare conflict (14%).

Overall, we only observed 13 cases (3%) in which the robot had to make an emergency stop because pedestrians did not yield. Among pedestrians who did yield, when they yielded—early on or last minute—seemed to be normally distributed, generally peaking on yielding somewhat early on. As one might expect, there was a very strong and clear pattern with pedestrians yielding earlier on getting much closer to the robot than pedestrians who yielded later (this averaged to 4.29 for the intimate zone, 3.05 for personal, 2.23 for social, and 1.58 for public). Yield timing did not change much with general busyness, with pedestrians only yielding a bit later while it was more busy than while it was averagely busy or less busy (average yield timing of 3.60, 3.17, and 3.21, respectively).

Conflict type For both from-behind and heads-on conflicts, in nearly all cases, pedestrians yielded by changing their trajectory (respectively 56/66 and 249/277), though quite a few also changed their speed (respectively 34 and 68); just changing speed was rare (respectively 8 and 9). For crossing conflicts, we saw that pedestrians would more often change their speed (41/51) than their trajectory (36/51). Pedestrians would, at times, also find a spot to stop and wait for the robot to pass.

The timing of the different yield behaviors did differ between conflict types (see Figure 3.4, top left), with pedestrians comparatively yielding later during crossing conflicts more often than during heads-on and from-behind conflicts.

Wideness We saw relations between several factors and the wideness of the passageway, be it narrow at $\approx 1.5m\%$ (narrow sidewalks/streets, 201 observations) or broader (broad sidewalk/street or open area, 225 observations). Overall, on average, pedestrians yielded only minimally earlier in narrow passageways than in broad ones (average yield timing 3.22 and 3.55, respectively). Yet when we look in-depth, more complex interactions become apparent when we look between a.o. group size/composition and age, more complex interactive frequency of the different conflict types also changed with wideness, as a more significant proportion of conflicts was heads-on for narrow (Crossing 12, Heads-on 176, From-behind 43) than for broad (Crossing 65, Heads-on 150, From-behind 55) passageways.

Group size There was a clear trend of group size influencing yield timing, with, on average, individuals yielding earlier than pairs yielding earlier than bigger groups (3 or more people) (Figure 3.4, top right). Groups yielded later in wide passageways than in narrow ones.

Gender Males (individual and in all-male groups) yielded marginally earlier on average than females (individual and in all-female groups) (Figure 3.4, bottom right). Mixed-gender groups overall had more intermediate yield timing than other groups and comparatively more often yielded earlier in narrow passageways and later (but not very late) in broad passageways.

Assumed age As shown in Figure 3.4, older pedestrians, on average, yielded later than younger pedestrians on narrow sidewalks but earlier on wide sidewalks. As age

3

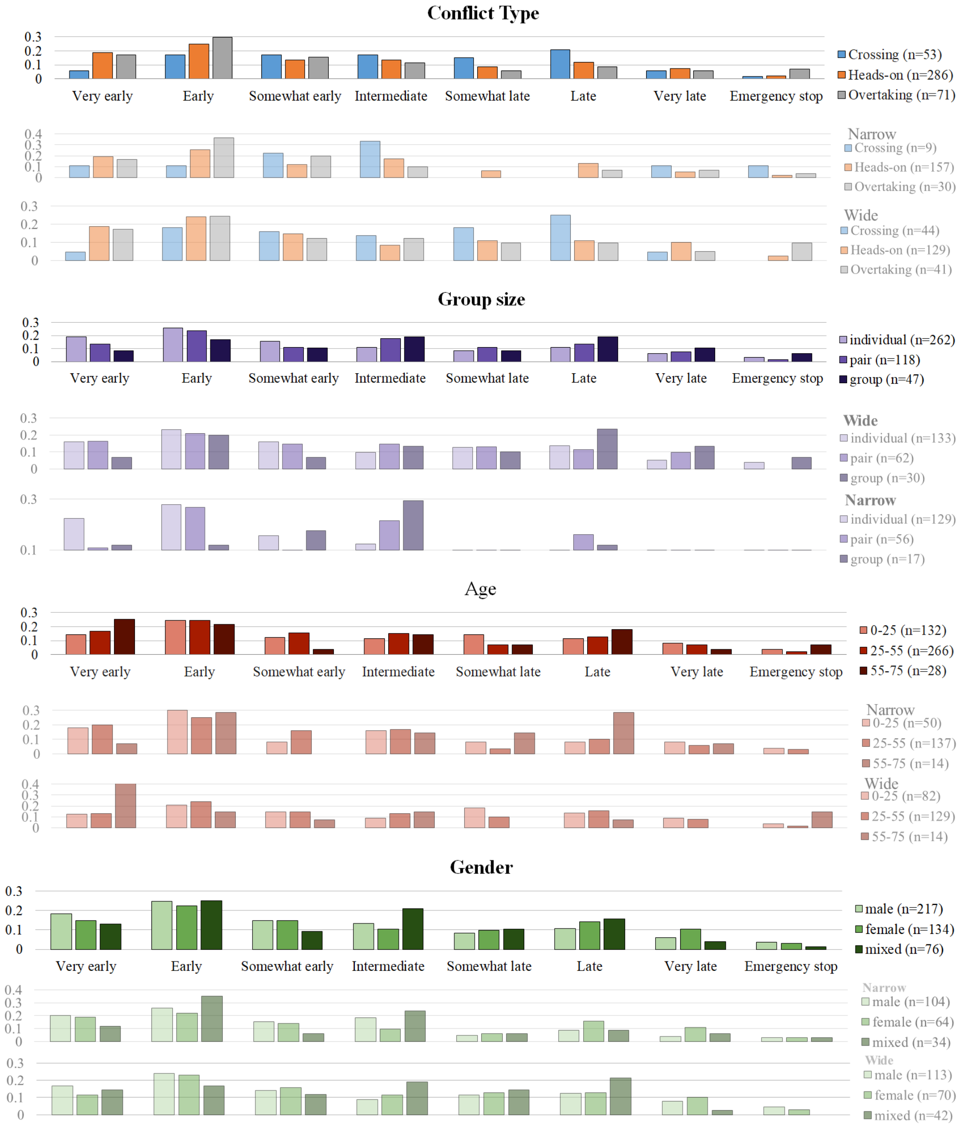


Figure 3.4: Distribution of different yield timings for different categories in various factors (conflict type, group size, assumed gender, and assumed age). Each factor is shown for all observations together (above) and split out on their effect in narrow and wide passageways (below).

was assumed by the observers, we checked this effect for a range of age groups, and the found pattern was consistent.

3.4.2. SOCIAL CUES AND PERCEPTION OF THE ROBOT

Within the subset of interviews (32 interviews involving 60 people, each conducted directly after an encounter), we found various factors that formed meaningful patterns with how strongly people felt the robot should change its behavior. To quantify these patterns, we coded their answers on the extent it indicated the robot should adapt its behavior according to at least one of the interviewees (5-point Likert scale, from 'not at all' to 'extremely'). After resolving bigger ambiguities with a third coder, our two coders strongly aligned (Pearson's rho of .96). We used average scores, rounded down, in our exploration of patterns. Though the evidence is limited, we share these informal descriptions of these patterns here as directions for future work.

Most of our interviewees were positive about the robot when asked how they experienced running into it–with recurring qualifiers being "strange", "interesting", and "nice". Only in 5 interviews did people indicate some discomfort at first. Yet, when we asked if the robot should adapt its behavior, people were more forward with their comments (11 interviews where at least one interviewee indicated the robot should change its behavior, seven interviews with not all, and 14 interviews in between).

On narrow streets, our interviewees were strongly skewed towards wanting the robot to adapt, while on broad streets, they were more likely not to want the robot to adapt.

Pedestrians strongly felt that the robot should adapt its behavior often gave the robot less space (skewed towards passing intimately close) and spent much longer looking at the robot before passing it (skewed towards looking at it 80-100% of the time). Pedestrians who changed their speed and/or did not change their trajectory were also comparatively more likely to strongly feel the robot should adapt its behavior (66% and 40%, relatively). Pedestrians who felt strongly that the robot should adapt, as well as those who felt it need not adapt, often yielded (very) early on (in 9/12 and 4/5 encounters, respectively), while those who somewhat felt the robot should adapt mostly yielded immediately fast (7/15 encounters) or later (4/15).

3.4.3. OTHER OBSERVATIONS

Throughout our observations, we noticed numerous other patterns. While the evidence for these patterns is more circumstantial (usually just a few observations) and thus would benefit from more extensive study, they were rich and relevant enough to justify their inclusion.

Stationary pedestrians Even though we only encountered 14 conflicts with stationary pedestrians, five resulted in an emergency stop. This accounts for 5/13 emergency stops. Even with the relatively small set of observations, this strongly suggests that stationary people are less willing to yield to a robot.

Engagement and teasing The robot attracted various kinds of attention, from kids wanting to sit on it (2x), to people stopping or slowing down to look (13x), film (4x) or take pictures (1x). Several pedestrians teased/tested[1] the robot by deliberately blocking its path (9x), jumping over it (2x), touching it (3x), or making sudden/kicking motions (3x). In two encounters, people gestured to/yelled at the robot–not counting the many instances where people talked *about* the robot.

In other words, the robot faced a very broad array of non-verbal reactions, many of which could be disruptive to its disrupting our observations; these added up to 9% (40/427) of all observed conflicts. While this number is likely to be partially due to a novelty effect, this still ought to be a factor of concern for any real-life deployment of such systems.

Interestingly, while some pedestrians blocking the robot resulted in emergency stops (3/9), in most of these encounters, they instead yielded (very) late (5/9). This aligns with the feeling our observers got that the lack of response caused by the protocol to ignore humans was what led people to back off in the end.

Objects; cycling, boxes, and strollers One factor we annotated was the presence of objects that might influence the interaction. The most common were people on bikes (71 conflicts), dogs (15 conflicts), people walking with a bike (7 people), and strollers (7 people). But this also included big boxes, luggage, garbage cans, and babies. For some of these objects, we saw patterns emerge.

While not a common occurrence (n=3), each time the robot ran into someone carrying big boxes, the robot operator noticed afterward that they had subconsciously broken protocol and yielded. This suggests, even with the evidence being only circumstantial, that pedestrians carrying objects should have a special status in social navigation—possibly to accommodate their line of sight being blocked.

Cyclists (n=71) clearly kept more distance from the robot. While this was not fully captured in our observation forms, the observers' impression was that they yielded further away in the distance to accommodate their higher speed.

People with strollers (n=7) kept less distance when yielding and often looked at the robot. To the observers, this was read as expecting the robot to yield and subsequently giving it less space while yielding as a social signal–in line with the discussed pattern that people who strongly felt the robot should change its behavior gave it less space.

3.5. Discussion

During our observations, we saw that most often, pedestrians yield for a robot and avoid potential conflicts long before collisions would have occurred. On the other hand, there were also cases in which pedestrians did not yield or did so at the last minute. Based on our observations, we identified various factors that correlate with this yielding behavior; an overview can be found in Table 3.1.

[1] To our observers, these behaviors sometimes felt like bordering on harassment/vandalism. Yet, whether they should be framed as such depends on the observed pedestrians' underlying motivations, such as curiosity. While we explored this in other (ongoing) work, the results of this chapter do not justify strong conclusions either way.

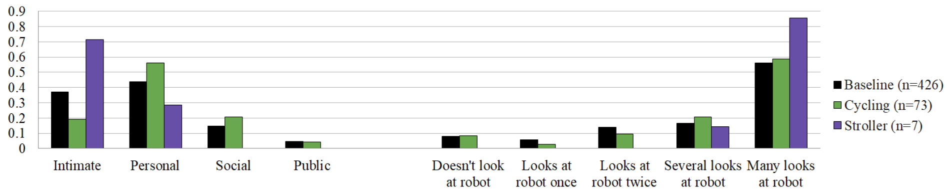


Figure 3.5: Distribution of relative frequency of reactions to the robot in terms of closest proxemic distance (left) and number of looks at the robot before passing (right) for pedestrians that are cycling and pedestrians with strollers, as compared against the baseline in all observations.

3

These factors show yield behavior and yield expectations to be a continuous complex coordination of who will/should yield (when and how). In this, one's yield behaviors simultaneously serve an implicit communicative function; not yielding (yet) indicates to the other that one is less likely to yield. This is akin to how proxemics behaviors are viewed in later frameworks like the intimacy equilibrium model [275–277] and dynamic implementations in human-robot interaction [266, 278].

From this understanding, we can cluster the factors found on their role in coordinating the yield behavior and identify design considerations for accommodating them in developing robot social navigation behaviors (Section 3.5.1). We then reflect on how the specific context of our field study may have influenced our findings to frame opportunities and limitations based on the generalizability of the work (Section 3.5.2). As such, our reflections on the method provide an additional contribution, enabling others to elicit these kinds of insights in similar contexts (Section 3.5.3).

3.5.1. DESIGN CONSIDERATIONS FOR COORDINATING YIELD BEHAVIOR

Empathic awareness First and foremost, the found factors allow us to highlight considerations for what should be understood of the situation in the design of social navigation. Crucial in this is the observation that the observed factors do not seem to be independent, from the pattern of yielding behaviors of older pedestrians changing with the wideness of the passageway to the observed relation between people, indicating the robot should adapt its behavior and their behavior in gaze, proxemics, and yield timing. Even though some of the factors we observed have already found their way into various works on socially aware robot navigation [279], as of yet, many of these factors and social cues are often studied in isolation and/or not as factors weighing into the interaction as a whole.

When interpreting these interdependencies between factors, a parsimonious explanation is that they weigh into two (possibly implicit/subconscious) social decisions involved in sidewalk navigation; this interpretation also ties into our earlier reframing of these interactions as a continuous complex coordination. Firstly, several factors can be tied to *awareness*—are people aware of the robot—including attention,

Factor	Observed pattern	Evidence
Conflict type	When crossing, pedestrians are more likely to yield early on by changing their speed. During head-on collisions, yielding by changing trajectory is most common.	Figure 3.4
Wideness	In addition to the patterns in combination with age and group size listed below, people overall also yielded somewhat earlier in narrow passageways.	Figure 3.4, our interviewees more strongly wanted the robot to adapt after encountering it on a narrow street
Stationary	When pedestrians were stationary, they were less likely to yield to the robot.	36% (5/14) of encounters with stationary pedestrians led to emergency stops
Group size	Individuals yielded earlier than pairs than bigger groups. Groups yielded later in wide passageways.	Figure 3.4
Gender	Males yielded marginally earlier for the robot than females, both individually and in groups.	Figure 3.4
Assumed age	Older pedestrians yielded later than younger pedestrians on narrow sidewalks but earlier on wide sidewalks.	Figure 3.4
Social cues	Whether a pedestrian perceived the robot's behavior as appropriate were reflected in their space given, gaze, change of speed but not trajectory, and yield timing.	32 follow-up interviews
Engagement & teasing	Pedestrians commonly directed actions towards the robot, even to the point of initiating conflicts, from trying to block the robot to playing with and jumping over it to stopping to film it.	Deliberate non-verbal actions towards robot in 9% (38/427) of conflicts.
Carried items	When running into someone with big boxes, the robot operator subconsciously yielded	100% (n=3) of operator subconsciously breaking protocol
Cyclists	Cyclists kept more distance	Figure 3.5 (n=73)
Strollers	People with strollers kept less distance and looked at it much longer, presumably expecting it to yield	Figure 3.5 (n=7)

Table 3.1: Overview of the found factors, the pattern, and the evidence for it.

gaze distractions by cell phones and people carrying items; even the static groups more often leading to emergency stops may, in part, be caused by these groups often being turned towards each other and thus having less awareness of what is going on around them. Secondly, several factors can be tied to *cost*—how much effort will yield take for the other and how to fairly weigh that effort—including factors such as strollers, age, the wideness of path, standing still, and group size. From this perspective, for example, our observation that older pedestrians tended to yield later on narrow sidewalks may be explained by assuming that some had somewhat reduced mobility and were expecting to be accommodated.

Following these reflections, we thus suggest interpreting the factors jointly to provide information for making social decisions related to awareness and cost in social navigation. Instead of responding to individual factors such as age and group size on their own, consider how awareness and cost. These factors influence awareness, and cost goes beyond surface-level understanding of the factors by themselves and enables what we will refer to as *empathic awareness*. While a few papers try to model pedestrians and their decision-making processes for robot social navigation [280], these do not include most of the factors we pinpointed.

In all, we thus recommend, from a design point of view, increasing such empathic awareness in social navigation, enabling social navigation that more deeply understands people. On a technical level, this will require bridging the gap between considering these factors and modeling their impact on the decision-making process. Moreover, reliable detection of many of these factors is likely to be non-trivial (e.g., [281]); our findings suggest some low-hanging fruit that may already cause key improvements when implemented, such as detecting when pedestrians are distracted by cell phones and detecting objects that limit maneuverability (e.g., strollers and carried items). Nonetheless, as this means a more full-fledged empathic awareness will likely remain unfeasible in the near future, designers and developers should further consider how (technical) boundaries of a robot's empathic awareness can be expressed and communicated.

PA responsiveness Secondly, in our observations and interviews, whether a pedestrian perceived the robot's behavior as appropriate was more clearly reflected in affective social cues like space given, gaze, and change of speed, than in more functional behaviors, such as trajectory and yield timing. Such affective, social cues have been extensively studied in social robotics [282, 283], yet work on socially aware robot navigation still commonly reduces humans to trajectories over time (as also concluded in [11]); though recent efforts have started to include more affective states (e.g. [284]). Our impression is that these affective cues play a necessary role in continuously communicating yield expectations—in line with our earlier reframing of yield behavior as mutual, continuous, complex coordination.

It suggests that responding to human's perceived appropriateness (PA) by reading and responding to effective cues—a more *PA responsiveness*—would provide a crucial opportunity for robots to improve their appropriateness. In line with earlier work on robot social positioning [266, 285, 286], what is appropriate may be something that is established through this back-and-forth of responding to each other's affective social cues (see also [287]).

From a design point of view, we would thus recommend considering how this kind of PA responsiveness might afford new kinds of interaction. While early detection of these signals may still be unreliable [288], there would already be a significant impact of being able to repair severely escalated situations in social navigation. Pending robots understanding such affective cues, designers could also deliberately create affordances for pedestrians to trigger responsive behaviors[2], empowering pedestrians in shaping how they want a robot to navigate around them.

Readable expressiveness Thirdly, as we greatly constrained how our robot acted around people, this also gave us clear insights into how that affected people's reactions, particularly when contrasted with less constrained observations. In numerous other explorations with more social behaviors, we have seen people get engaged with such robots, teasing them, and even getting to the point of vandalizing them (e.g., [240]), yet, in this chapter, while people would still approach and 'test' the robot, its lack of social response seemed to be an effective approach in having them get out of the way for it[3]. So even if a behavior may 'just' be intended to ensure safety or even defuse teasing, due to how it is read, it may still play out differently—even to the opposite effect.

These findings suggest that people will continuously read and respond to the behaviors of the robot, implying that behavior design should be mindful of how the behaviors of the robot will/can be read by pedestrians: *readable expressiveness*. One aspect of this that has been extensively studied is to deliberately design behaviors to provide information about the aims and intentions of the robot—i.e., legibility [265, 289]. Beyond this, another crucial aspect will be to consider how pedestrians may read the (combined) behavior robot's of a robot and assume beyond what was intended, i.e., projecting their expectations on it.

For design, we would thus recommend the express design of behaviors that convey what is intended and, crucially, as little else as possible. This can likely build on earlier work attempting to make robots intuitively relatable to people, e.g. by leveraging anthropomorphism[4]. Here, in line with the set-up and observations in this chapter, minimal behaviors may be more effective in helping avoid conveying unintended additional meaning, such as minimum channels for maximum tuning. Crucially, this also hints that designers should avoid the risk of over-anthropomorphising.

3.5.2. Generalisability

Within our observations, the various factors are jointly related to yield behaviors. The most notable example is how the patterns for age and group size were different depending on the wideness of the passageway. We would thus argue that the factors

[2]For example, what if a robot had a red area with a footprint on it that pedestrians could use to push it away with their feet if they were upset with what it was doing? Or buttons you could press to rate its performance?

[3]Admittedly, in two separate instances people moved out of the way by jumping over it, which would still be undesirable as it risks harm to robot and pedestrian

[4]Given the focus of our work on movement in interaction, with a robot that has little in shape that readily relates to the human body, calling it *anthropomorphism* may be more appropriate.

themselves should thus be seen as context-dependent; for example, whether or not people are aware of and willing to yield for a robot may well be different between a packed market street with scooters zipping by and a potholed sidewalk beside a road.

We want to highlight in particular the context factors of time and novelty; it is very likely that when robots become more commonplace in our public spaces, our expectations and reactions will also change. Depending on how these robots are introduced and behave, people may gradually become more understanding of their needs and more/less willing to accommodate them (e.g., [243]).

Furthermore, it is worth emphasizing that we conducted this chapter at one specific location and thus in a relatively narrow context. Among others, we suspect that the relatively car-free nature of the area may have increased participants' willingness to step onto the 'road' in yielding for the robot; a few interviewees asked if the robot was for bomb disposal, an association likely based in our location being a former navy base; the terrain being commonly used for leisure may have also skewed our results.

In addition, various other aspects of robot social navigation may be relevant but have not been the focus of this chapter. For example, we noticed how people (and dogs) would often be intrigued by the robot but only turn to look at and comment on it after the conflict was resolved—in line with the findings in [272]. We also noticed that the built environment, e.g., road obstacles and raised curbs, seemed to influence the yielding behavior of people; this might be a worthwhile direction for further study. We did not see many vulnerable pedestrians in our sample, though they should also be a factor to consider [239]. Implicit biases and other limitations of our observers may have led to factors being missed, though this does not invalidate the factors found.

Most crucially, though, the sheer amount of factors we identified suggests that, at the very least, it would be relevant to do similar confrontations in other contexts, e.g., as such robots become more commonplace. The comparisons with other works we discussed throughout this chapter suggest that the factors we found are likely to also apply to other contexts. Beyond that, both the method outlined in this chapter and the framework provided by our design considerations can contribute a valuable starting point to designers and developers alike.

3.5.3. CONTEXT CONFRONTATION; LESSONS AND RECOMMENDATIONS

The efforts outlined in this chapter align with previous studies and their findings. Our finding on robot teasing aligns with the aforementioned prior work [240, 242, 249], but our 'solution' of the robot simply not yielding is new. Our findings that conflicts are often (but not always) resolved easily by pedestrians much align with the earlier fieldwork [253, 272, 273]—even though our robot seemed to be less attractive to dogs [272].

In contrast to this earlier work, we added more systematic evidence by including our observation forms. These allowed us to quantify our observations better, in terms of fry, without requiring us to record the observed pedestrians and infringe on their privacy. Doing so well required us to overcome a few hurdles in designing the study:

(1) We found that observers needed a training session to be explicitly aligned on the more subjective judgment calls involved, (2) we had to cap the detail in annotating the cues to stay within the bounds of what our observers could feasibly recall of encounters, which limited the reliability of those observations, and (3) observers need practice in staying close enough to the robot to make detailed annotations, but far away enough not to be immediately associated with it. Throughout the process, we repeatedly noticed that many of the aspects noted on the observation form were very easy to overlook/forget when not reminded about them by the observation form, presumably because their influence is usually more subconscious. Using the forms as a framework provided the support needed while allowing unanticipated observations to be made. In all, the observation forms were essential in providing extra evidence and informing our considerations, helping us retroactively tease out patterns and factors from the complex interplay of factors that yield behavior turned out to be.

Furthermore, we deliberately conducted our observations on a robot that displayed no social behaviors. While this robbed us of the opportunity to pilot such behaviors and explore their effectiveness, it did provide a uniquely 'clean' interaction, which we found strongly supported us in identifying the different factors more clearly. It was also instrumental in our reframing of yield behavior as a continuous complex coordination—a negotiation, a dance—which enabled our design considerations.

3.6. Conclusions

We conducted a field observation study during which we tele-operated a robot to navigate in an open public space while explicitly ignoring humans. This resulted in 427 encounters, which we documented using a structured observation form to identify factors influencing human-yielding behavior. Additionally, we conducted a small sample of 32 interviews to explore the relationship between human perceptions of the robot's behavior and social cues such as proxemics, gaze, and facial expressions, further deepening our understanding.

Our findings reveal that human yielding behavior is influenced by several seemly dependent factors, including human activities (cycling, carrying items, running) and demographics (age, gender), the robot's behavior, and environmental settings (width of passage). These factors might tie into two social decisions: **awareness** and **cost**. Therefore, we suggest considering how these factors influence awareness and costs **empathic awareness** that aims to enable social navigation that more deeply understands people.

Secondly, observations and interviews indicated that pedestrians' perceived appropriateness of the robot's behavior was more evident in affective, social cues—such as proxemics, gaze, and change of speed than in functional behaviors like trajectory and yield timing. Therefore, reading and responding to the perceived appropriateness (PA) through those cues—a more **PA responsiveness**—would provide a crucial opportunity for robots to improve their appropriateness. We thus recommend considering how this kind of affective responsiveness might afford new kinds of interaction.

The study reveals that pedestrian interpretations of robot behaviors can deviate from their intended safety or conflict resolution functions. This variation highlights the need for designing robot behaviors with **readable expressiveness**, ensuring that they communicate tended actions while avoiding misinterpretation.

Overall, this chapter identifies various factors contributing to the yielding behavior of pedestrians, particularly human activities and demographics, robot behavior, and environmental settings. It also highlights the significance of human social cues and emotional responses to human perceived appropriateness of robot behavior. These insights enrich our understanding of the dynamics of human-robot interactions and pave the way for further investigations into detecting and adapting inappropriate robot navigation, ultimately enhancing human-robot interaction quality in social environments.

4

PARSNiP Dataset for Perceived Appropriateness Detection

Despite advancements in socially aware navigation, robots still often behave inappropriately in social environments. To ensure their successful application, robots must detect the human perceived appropriateness of their navigation behaviors. More importantly, such inappropriate behaviors largely happen in narrow environments. Therefore, it is important to answer the research question: How can we detect the perceived appropriateness of robot navigation behavior in narrow environments (RQ3)? We present a novel dataset that covers a complete range of perceived appropriateness and uniquely incorporates human emotion and attention to facilitate the perceived appropriateness detection of robot social navigation in pathways (PARSNiP). It is created based on a series of human-robot interaction experiments with 30 participants and a mobile robot. Several typical machine learning models are utilized to evaluate the dataset and analyze the contributions of different features in detecting perceived appropriateness. The results indicate that incorporating emotional and attentional features can significantly improve the accuracy of perceived appropriateness detection. There was an increase from 63% to 68% using algorithm-predicted emotional and attentional features and a further increase to 79% with the emotion and attention data reported by the participants. With the dataset, researchers could train machine learning models to enable robots to detect perceived appropriateness accurately, fostering adaptations that improve their responsiveness and accuracy in social interactions.

This chapter has been submitted as: Yunzhong, Z., Vroon, J., Zoltan, Z., Gerd, K., 2024, PARSNiP: A Novel Dataset for Better Perceived Appropriateness Detection in Robot Social Navigation using Emotion and Attention Features.

4.1. Introduction

Mobile robots are increasingly operating in social environments such as offices, hospitals, and public spaces, where they encounter and interact with humans. These robots undertake different tasks, including assisting in nursing and patient rehabilitation, delivering parcels to homes and offices, stocking shelves in warehouses, and cleaning floors in shopping malls [290]. A fundamental prerequisite for the successful applications of mobile robots is their ability to navigate around humans [291]. Robots that fail to respect humans and behave inappropriately have been restricted or banned in social environments [292].

Socially-aware navigation, also referred to as human-aware navigation [6], socially compliant navigation [11], socially acceptable navigation [8], or socially competent navigation [9], aims to enable robots to navigate social environments safely, effectively, and in a socially acceptable manner. Its primary objective is to integrate social norms into robot navigation behaviors [293], making them, for example, respect personal space [120], minimize path interference, and prioritize human behavior [156]. However, pre-defined social norms lack flexibility and fail to consider contextual factors and individual differences adequately. More sophisticated algorithms have been developed to enable robots to detect and respond to social dynamics, including human physical activities [294], and social signals such as emotion [42, 295], intention [296, 297], and dominance [52] in varied settings including crowded scenes [172], narrow spaces [298], and urban environments [299].

Current studies in socially aware navigation primarily focus on robots perceiving and responding to humans [290, 300], yet they often overlook how humans perceive robots, especially real-time perception. Just as humans pick up on social cues and adapt to others' feedback, robots should also detect whether their behavior is perceived as appropriate by nearby individuals. Humans use a rich set of social cues to indicate if they perceive behavior as inappropriate, annoying, or unsafe, such as facial expressions, hand gestures, and evasive motions. Due to the limited studies and datasets on the perceived appropriateness (PA) of robot navigation behavior, current robots are unable to detect the PA of their navigation from these cues.

We introduce the PARSNiP dataset, which contains PA labels and uniquely includes emotion and attention, to improve robots' ability to detect PA. Collected from human-robot interaction experiments that involve 30 participants and a robot, PARSNiP offers a complete range of PA levels and is enriched with features crucial for PA detection, such as the motion features of both humans and the robot, as well as human emotion and attention. Several typical machine learning models are employed to analyze the dataset, and findings reveal that incorporating emotional and attentional features markedly improves PA detection performance. This dataset can be utilized to develop machine learning models, which could then be applied in robots to enhance their ability to accurately detect PA and adapt to human behaviors more effectively.

The remainder of this paper is structured as follows: Section 4.2 reviews existing datasets and their limitations, which motivates the creation of PARSNiP. Section 4.3 describes the dataset creation, including collection and processing methods. Section 4.4 presents the testing of the dataset on multiple machine learning models, with an analysis and discussion of the experimental results in detecting PA. Finally, Section 4.5 makes conclusions and outlines future directions.

4.2. Related work

Many datasets have been developed to improve and evaluate socially aware navigation systems. These datasets primarily focus on gathering human motion data to assist in robot imitation learning [301], covering a variety of environments, both indoors [302–308] and outdoors [21, 308–315]. These datasets are used to train data-driven methods, such as reinforcement learning [102] and deep learning techniques [316], for simulating robot social navigation behaviors. However, they do not take into account the decision-making processes and psychological states of humans.

Additional datasets have integrated affective features, extracted from posture and movement cues [317], to enable the detection of human emotion [165]. Although these datasets have enhanced robots' ability to navigate socially in human environments, they often fail to consider feedback regarding human's perceived appropriateness of the robots' navigation behavior. In response, recent datasets have shifted focus to social errors or mistakes, enabling robots to recognize and correct their inappropriate behaviors [318, 319]. While these studies are not within the navigation context, they have yielded significant insights. They demonstrate that social cues play a crucial role in detecting the perceived appropriateness of robot behavior [62, 64], and indicate that integrating emotion and attention could further enhance PA detection performance [318].

In robot social navigation, only 1 dataset considered and enabled PA detection, albeit with limitations such as focusing solely on robot inappropriate positioning behavior and including only low-level motion features [10]. We create the PARSNiP dataset to offer a complete range of PA levels and uniquely incorporate emotion and attention to enhance PA detection.

4.3. Dataset creation

We make the assumption that human and robot motion features, together with human emotion (through valence and arousal, where valence refers to how positive or negative an event is, and arousal reflects whether an event is exciting or calming [320]) and attention, contribute to the detection of perceived appropriateness (PA) of robot navigation behavior, as shown in Figure 4.1. Emotion and attention are intermediate features that can be predicted from low-level human and robot motion features [163, 321].

The workflow to create the PARSNiP dataset is shown in Figure 4.2. It consists of 4 major steps: Step 1 involves the human-robot interactions, with the participants interacting with a robot 8 times. Step 2 is the data collection, including RGB-D recordings and robot odometry to capture human and robot motion, along with questionnaires to gather participants' subjective annotations of emotion, attention, and PA. In Step 3, the data processing, human and robot motion features are

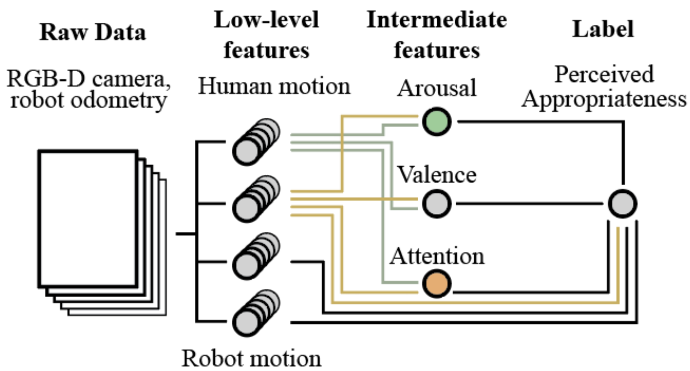


Figure 4.1: Assumption for PA detection.

extracted and computed from the raw data. Step 4 is the dataset building, which integrates the computed human and robot features with attention, emotion, and PA to build the PARSNiP. The experiment was approved by the ethical committee of the university.

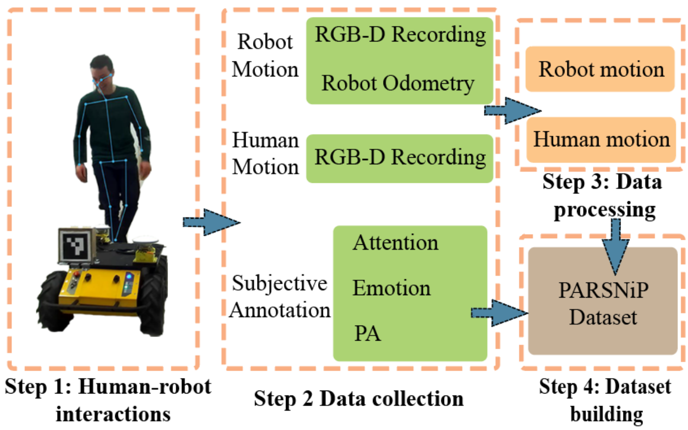


Figure 4.2: The workflow for creating PARSNiP.

4.3.1. Pilot study

A pilot study involving both field observations and controlled human-robot interactions has been conducted to determine the setting and types of interactions that contribute to triggering a complete range of PA.

The field observations involved a mobile robot (Clearpath Husky) navigating in diverse outdoor public spaces: narrow pathways, open areas, cross-shaped roads, and street corners. A robot operator fully controlled the robot's navigation behavior to observe the robot's behavioral impact on humans in a natural but safe way.

Narrow spaces consistently elicited richer and more nuanced reactions and PA of the robot [322], consistent with the literature indicating a higher incidence of human-robot conflicts in narrow spaces [80]. A subsequent controlled study involved 8 participants interacting with the robot in pathways of different widths (1.0m, 1.2m, and 1.4m). Based on observations and questionnaires, a pathway width of 1.2m was selected, striking a balance between the richness of PA and human safety. Importantly, the richness of PA persisted even when the pathway was merely marked by floor indicators. During the controlled study, interactions that were highly effective in eliciting diverse human reactions and PA were identified, such as the robot making sudden changes in trajectory and velocity near humans and unconventional human-robot spatial relationships like blocking or squeezing paths.

4.3.2. STEP1: HUMAN-ROBOT INTERACTIONS

Setting The setting is depicted in Figure 5.1, which is a pathway delineated by tape markers on the floor. An RGB-D ZED2 camera, positioned on a table 0.75m high and 1m away from the end of the pathway, is used to capture the participants' full-body motion. The camera is oriented towards the participants' walking direction to record videos (1080p, 30fps). This setup was chosen to ensure the collection of clean, comprehensive data on human-robot interactions, avoiding the limitations of an onboard camera that has a restricted field of view or experiences data occlusion when humans are in close proximity.

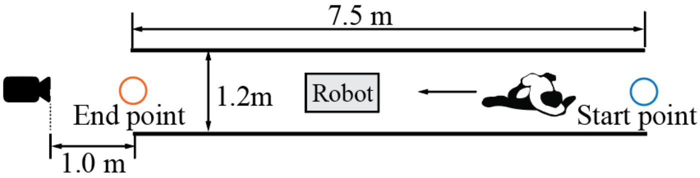


Figure 4.3: Overview of the setting. Each participant walks from the start point to the end point for each interaction.

Designed interactions 8 interactions are designed to trigger a complete range of PA of robot navigation behavior. This is built upon the study of Koay et al., which identified specific interactions causing human discomfort [323], as well as insights from the pilot study. Husky A200 is used as the robot platform to interact with the participants (dimensions of $0.90 \times 0.67 \times 0.39\,\mathrm{m}$) [324]. Each interaction with the robot and human start and end points is visualized in Figure 4.4 and detailed below:

- **Block (Interaction 1)**: The robot moves toward its goal and positions itself to block the participant's path at a close distance, potentially triggering feelings of intrusion or frustration as it forces the individual to stop or change direction.
- **Change Direction (Interactions 2, 3)**: The robot moves toward its goal and changes its direction to approach the participant from different angles at a

close distance, expecting the participant to adjust their trajectory to avoid collisions.

- **Squeeze (Interaction 4)**: The robot moves toward its goal while squeezing the participant's path, which may be perceived as invasive and cause the participant to feel pressured.
- **Stop (Interactions 5, 6)**: The robot moves toward its goal, and makes sudden stops in front of the participant at a close distance, potentially interpreted as an interruption, forcing the participant to stop or navigate around the robot.
- **Accelerate (Interaction 7)**: The robot moves toward its goal and unexpectedly accelerates, which could alarm the participants and may be perceived as aggressive or unsafe.
- **Stationary (Interaction 8)**: The robot remains stationary within the environment, serving as a control condition expected to be socially acceptable and minimally intrusive the whole time, providing a baseline for the participants' perceptions.

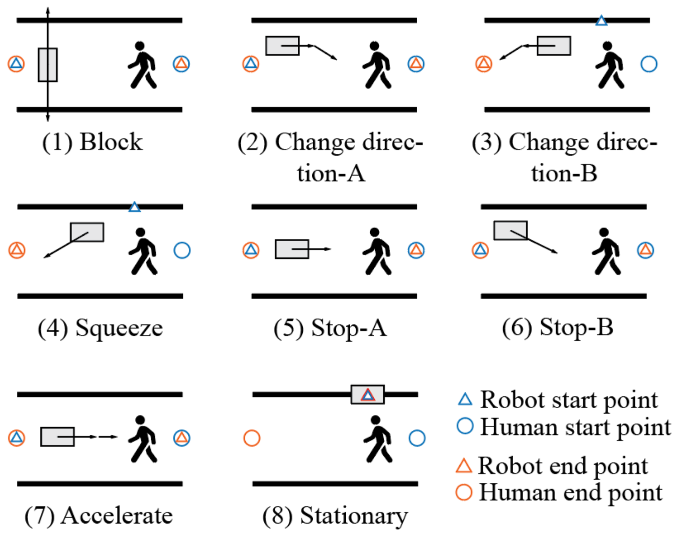


Figure 4.4: The sketch of robot-human directional relationships for 8 interactions. During each interaction, the PA of the robot's behavior may dynamically vary based on factors such as distance, speed, acceleration rate, and other variables.

Many of these interactions involve many appropriate robot behaviors, such as maintaining a distance greater than humans' public space [325], with the robot passing by and then moving further away from humans. Additionally, these interactions also include the robot exhibiting inappropriate behaviors, particularly when it is at a close distance from humans. As such, these interactions enable the collection of a complete range of PA levels.

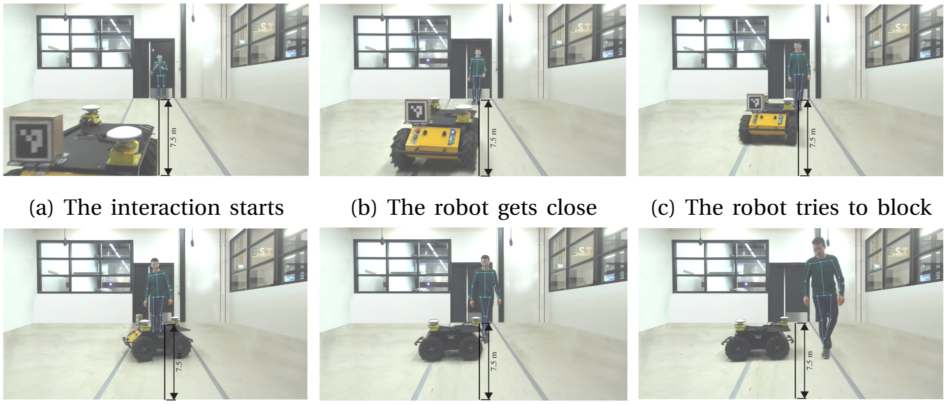
Experiment Before the experiment, all participants are given informed consent and required to walk around to familiarize themselves with the setting. Thereafter, each participant traverses a distance of 7.5 m from the starting point to the endpoint in the pathway, interacting with the robot 8 times in a random order (see Figure 4.4). This randomization helps prevent any order effects that could bias participants' responses or learning effects that could influence their PA of the robot's behavior over time. They have the freedom to navigate this distance as they see fit without additional guidelines. The robot typically starts from the endpoint of the pathway, except for interaction 8, where it remains stationary, and for interactions 3-4, where it is positioned 1.5m ahead of the participants to trigger specific interactions. All relevant positions are clearly marked to ensure that both the participant and the robot start from the same location in each trial, ensuring consistency across trials. An experienced operator was trained to manually control the robot during the interactions. Utilizing the Wizard-of-Oz method, the operator simulates autonomous navigation, safely initiating the intended interactions [326]. During the interactions, both the robot and the participants could step outside the pathway for safety or other concerns, similar to what might occur in real-life pathways. Figure 4.5 depicts real-life human-robot interactions, with the detected skeletal points visualized.

4

Demographics 30 participants were recruited through onsite convenience sampling, including 11 females and 19 males, with no exclusions. Most participants (27) were young adults between 18 and 34 years old, while 3 were adults aged 35 to 54 years. The majority (26 out of 30) were Dutch. Regarding robot experience, most participants had limited interactions with robots: 23 had no prior interactions with robots, 6 had some interactions, and only 1 had frequent interactions.

4.3.3. STEP 2: DATA COLLECTION

Upon completing all 8 interactions, the participants are presented with definitions of valence and arousal [320]. Afterwards, they are given questionnaires to indicate their current levels of valence and arousal. They are also requested to provide explanations for their responses to ensure that they understand how to assess and report their valence and arousal levels accurately. Meanwhile, the recorded interactions from the camera are divided into 2-second clips, resulting in 25-42 clips [327, 328]. These clips are presented to participants with online questionnaires to self-report PA, emotion, and attention, as data points of the dataset. PA is assessed on a 7-point scale ranging from -3 to 3, with participants asked to indicate the appropriateness of the robot's behavior (7-point scale: "The robot behaved completely inappropriately"–"Not sure / Neutral."–"The robot behaved completely appropriately.") [10]. Emotion is evaluated in terms of valence and arousal; both are also measured on a 7-point scale from -3 to 3. Attention is rated on a 6-point scale ranging from 0% to 100%. Additionally,



(a) The interaction starts (b) The robot gets close (c) The robot tries to block

(d) The robot blocks the path (e) The human tries to pass (f) The human passes

Figure 4.5: Human-robot interactions, with the human's 18 skeletal points captured by the ZED SDK.

participants are asked multiple-choice questions aimed at identifying perceptions of robot behaviors concerning PA value below 3, selecting from a list of factors identified in the pilot study. These factors include moving too fast or slow, taking sharp turns, getting too close, moving unnaturally, being too noisy, blocking the path, appearing threatening, squeezing the walkable area, making sudden changes, and an “other" option.

4.3.4. STEP 3 &4: DATA PROCESSING AND BUILDING

After the interactions and the questionnaires, all data points are collected. For each data point, the raw data from recorded interactions and robot odometry undergo processing to extract human and robot motion features. The ZED Software Development Kit (SDK) is utilized to extract human motion data, comprising 3D positions of 18 skeletal points and the body center for each frame. A 95% detection confidence threshold is established using the “Accurate" detection model, which fills in missing keypoints using historical data and human kinematics. The robot motion data are captured by the robot's odometer and an Aruco Marker, with the former providing consistent tracking and the latter ensuring reliable position data. The robot's data is transformed into the same coordinate system as the human data to ensure alignment for analysis. Using the Genetic Algorithm (GA) Toolkit in Matlab, this transformation minimizes the mean squared error compared to ground truth measurements and ensures consistent robot positional data for feature computation.

Based on robot and motion data processed, human and robot motion features are computed. Both the robot and human motion features are characterized by 5 statistical measures—minimum, maximum, mean, standard deviation, and min-to-max ratio—across each data point. Each point includes 30 robot motion features, including robot speed, acceleration, jerk, robot-human distance, robot-human direction, and robot turning angle [10]. Specifically, the robot-human approach angle

Table 4.1: Overview of the PARSNiP dataset. The dataset includes 922 data points collected from 30 participants, each interacting 8 times with the robot

Data type	Data (source)	Measurement (per 2-second interval)
Low-level features	Human motion (3D position of 18 skeletal points from RGB-D recordings)	11 categories of 360 human motion features, described by the min, max, mean, standard deviation, and min-to-max ratio: • Kinematic features (velocity, acceleration, and jerk (smoothed by low-pass filtering [329]) for head, wrists, and ankles) [321, 330], • Curvature for head, wrists, and ankles [321], • Quantity of motion (aggregated speed over a set of joints) for the arm and head region, upper body, lower body [321], • Bounding volume for the arm region, head region, upper body, lower body [321], • Displacement of joints for the head, wrists, and ankles [321], • Motion features (verticality, extension, left and right elbow flexion, left and right arm shape, hand relationship, and feet relationship) [321, 330], • Effort component of the Laban Movement Analysis (time, weight, flow) for the arm region, head region, upper body, lower body, and whole body [321, 330], • Spatial extension [321, 331], • Symmetry of the movement (horizontal, vertical, and bounding triangle-related symmetry) [321, 331], • Body balance (balance, center of mass for the arm region, head region, upper body, lower body, and whole body) [321], • Human gaze [332]
	Robot motion (Robot odometry, RGB-D recordings)	30 Robot motion features, described by the min, max, mean, standard deviation, and min-to-max ratio: robot speed, robot acceleration, robot jerk, robot-human distance, robot-human approach angle, robot turning angle [10]
Intermediate features	Emotion and attention (Self-report)	Attention (0% to 100%, 6-point scale), valence and arousal (-3 to 3, 7-point scale)
Label	PA (Self-report)	PA (from -3 to 3, 7-point scale)

4

is computed as the angle between the robot's current heading and the straight-line vector that connects the robot's current position to the human's position. These features could provide insights into the dynamics of human-robot interactions. For example, features such as robot-human distance and robot-human approach angle can reveal whether the robot is heading toward a person or will merely pass by. A small approach angle, combined with a low min-to-max distance ratio and a short mean distance, generally indicates that the robot is heading directly toward a person. Conversely, the same small approach angle, when paired with a high min-to-max distance ratio and a longer mean distance, might suggest that the robot, initially aimed toward the person, is moving away or adjusting its path. Additionally, a large approach angle with a low min-to-max distance ratio and a short mean distance could imply that, despite not facing the person directly, the robot is moving toward them.

Each data point also includes 360 human motion features, with 11 groups encompassing kinematic features, curvature, quantity of motion, bounding volume, displacement of joints, motion features, effort component of the Laban Movement Analysis, spatial extension, symmetry of the movement, body balance, and gaze [321, 330, 331]. Specifically, jerk derived directly from position data can be quite noisy. To address the inherent noise issues associated with this calculation, a low-pass filter is applied to the acceleration data [329]. The effort of the Laban Movement Analysis is computed according to the following functions [321, 330]:

$$E(t_i) = \sum_{k\in K} E_k(t_i) = \sum_{k\in K} \alpha_k v^k(t_i)^2 \tag{4.1}$$

$$\text{Weight}(t_i) = \max_{i\in[1,T]} E(t_i), \quad i = 1,2,3,\dots,N \tag{4.2}$$

$$\text{Time}^k(t_i) = \frac{1}{T}\sum_{i=1}^{T} \alpha^k(t_i) \tag{4.3}$$

$$\text{Flow}^k(t_i) = \frac{1}{T}\sum_{i=1}^{T} j^k(t_i) \tag{4.4}$$

By organizing human and robot motion features and human emotion (valence and arousal), attention, and PA for each data point, the dataset is thus built. A complete overview of the dataset is presented in Table 4.1.

4.4. Dataset analysis

4.4.1. Dataset distributions

Table 4.2 provides a comparison between the PARSNiP dataset and existing datasets in robot social navigation, highlighting its unique attributes and contributions to the field. Unlike its predecessors, PARSNiP encompasses a broad spectrum of human and robot features. It includes human PA, head, body, and hand motions, trajectories, emotion, and attention. The inclusion of emotion and attention marks a significant advancement, offering insights into better PA detection.

Table 4.2: Datasets for socially aware robot navigation

Dataset	PA	Head motion	Body motion	Hand motion	Trajectories	Emotion	Attention
UCY [312]	×	×	×	×	✓	×	✓
ETH [309]	×	×	×	×	✓	×	×
KITTI [313]	×	×	×	×	✓	×	×
Edinburg [310]	×	×	×	×	✓	×	×
Town [314]	×	✓	✓	×	✓	×	×
CFF [21]	×	×	×	×	✓	×	×
WildTrack [315]	×	×	×	×	✓	×	×
SCAND[311]	×	×	×	×	✓	×	×
VIRAT [303]	×	×	×	×	✓	×	×
ATC [302]	×	✓	✓	×	✓	×	×
L-CAS [306]	×	×	×	×	✓	×	×
KTH [304]	×	✓	×	×	✓	×	×
THOR [305]	×	✓	✓	×	✓	✓	×
EgoMotion [307]	×	✓	✓	×	✓	×	×
JackRabbot [308]	×	×	×	×	✓	×	×
UNC [317]	×	×	×	×	✓	×	×
Vroon [10]	✓	✓	✓	×	✓	×	×
PARSNiP	✓	✓	✓	✓	✓	✓	✓

PA DISTRIBUTION

Participants exhibited variations in their responses to the robot: 3 participants showed signs of disorientation and discomfort in the presence of the robot, even when it remained stationary. Conversely, 1 participant maintained a positive attitude towards the robot, showing high tolerance regardless of the robot's behavior. The range of observed human responses to specific robot behaviors varied significantly, especially during potential conflicts with the robot: Some participants paused to assess the situation before proceeding, while others subtly changed their path or quickened their pace to avoid the robot. A few participants turned their bodies to navigate past the robot.

Figure 4.6 shows the distributions of different PA levels, illustrating a complete range of PA, from -3 (completely inappropriate) to 3 (completely appropriate). The data show that the participants largely report the robot's behavior as completely appropriate, suggesting a somewhat imbalanced dataset, which highlights the need for further processing when using machine learning techniques.

Figure 4.7 illustrates the distribution of PA values across the 8 interaction types. Many interactions received a score of 0 to 2, indicating a neutral to slightly positive perception of the robot. However, distinct deviations are observed in interactions 1, 6, 7, 8. Interaction 1 of blockage shows a slightly negative value in median PA, while interactions 8 receive a positive peak, indicating higher PA. These variations in PA scores highlight how different types of interactions are perceived in terms of appropriateness, providing valuable insights into the contextual factors that may influence perceptions and the richness of PA levels.

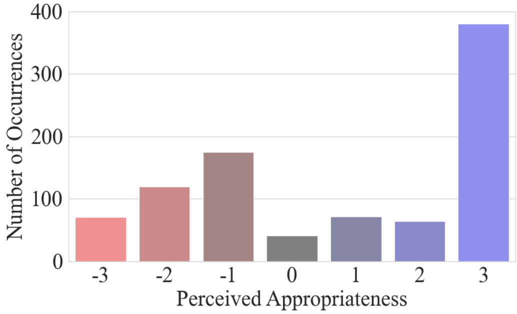


Figure 4.6: The distribution of different PA levels.

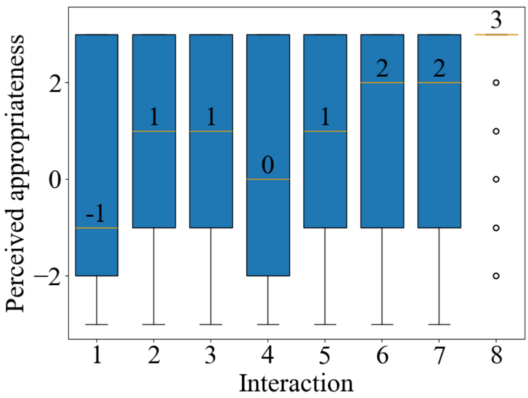


Figure 4.7: Interactions and PA levels.

The dataset includes 11 types of PA, detailed in Figure 4.8. Notably, “Block" with 173 instances, “Squeeze" with 112 instances, and “Sudden" with 110 instances are significant factors affecting the human perception of robot behaviors. In contrast, perceptions of “Threat", “Unnatural," and “Fast" are reported less frequently. Different interactions reveal variations in PA, providing insights into human-robot interaction dynamics. Interestingly, although interaction 8 of the robot being stationary generally receives positive evaluations, previous inappropriate interactions could potentially influence the perceptions of the robot. Overall, these 8 interactions contribute to a complete range of PA levels and richness of PA types, thereby enabling PA detection.

Emotion and Attention Distribution

Figure 4.9 illustrates the emotional responses and levels of attention across 8 different interactions.

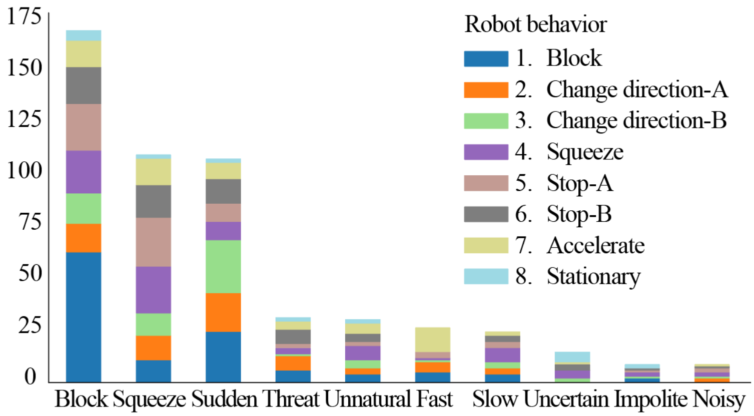
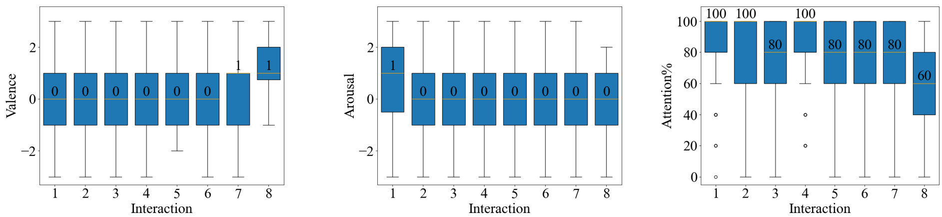


Figure 4.8: Interactions and PA types.



(a) Valence (b) Arousal (c) Attention

Figure 4.9: Valence, arousal, and attention distributions under different interactions.

Figure 4.9a presents the distribution of emotional valence across 8 interaction types. The median valence for interactions 1 to 6 remains consistently neutral, underscoring a uniform emotional response among these categories. Interactions 7 and 8 receive a higher median valence, indicating a more positive emotional reaction. Specifically, interaction 8 shows a narrower interquartile range, suggesting less variability and a more consistent response among participants.

Figure 4.9b illustrates the arousal levels across the 8 types of interactions. Interactions 2 to 8 display medians at zero, indicating no substantial deviation from a baseline arousal state. However, interaction 1 shows a higher median arousal level, which suggests that path blockage makes participants more excited compared to others.

Figure 4.9c shows that the attention metrics vary significantly across different types of interactions. Interactions 1, 2, and 4 achieve the highest median attention scores of 100%, reflecting heightened cognitive engagement or focus. Interaction 8 of the robot being stationary notably receives the lowest median and an expanded range, indicating fluctuating levels of participant attention and potentially lower engagement.

The impact of a robot's presence within the participant's field of view on human attention and emotion (valence and arousal) is also examined. Given the dynamic nature of human-robot interactions captured in 2-second intervals, direct comparisons are challenging. Factors such as the robot's movements into or out of view or transient presence complicate the analysis. Therefore, we differentiated the data points based on two scenarios: A. the robot is entirely outside the participant's view the whole time, and B. the robot has been within view.

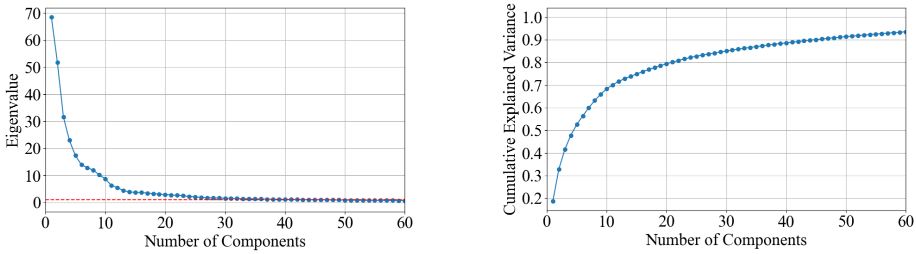
A Mann-Whitney U test was conducted comparing the differences between the two scenarios [333]. The statistical analysis revealed significant differences in attention and emotional responses. For Scenario A, the average attention was 82.7%, compared to 58.0% for Scenario B, with a significant p-value of 1.26×10^{-29}. This suggests more human attention when the robot is in view. Emotional responses also varied; valence averaged 0.73 in Scenario A and 0.18 in Scenario B, with a p-value of 5.5×10^{-6}, indicating more negative emotions when the robot is in view. Arousal was -0.44 for Scenario A and 0.25 for Scenario B, with a p-value of 6.0×10^{-8}, indicating that the presence of the robot within the view was more emotionally exciting for the participants.

4.4.2. DATASET ANALYSIS

CLUSTERING AND ALGORITHMS

A principal component analysis (PCA) is conducted to reduce dimensions and mitigate noise. This step entails calculating eigenvalues and the cumulative explained variances, as depicted in Figure 4.10. We select components with eigenvalues exceeding 1 [334], with 45 principal components that account for around 90% of the total variance.

The PA levels are binarized for further analysis, ensuring that the smaller group constitutes at least 40% of the data to enable balanced clusters. This leads to 2 distinct clustering strategies: Strategy 1 assigns PA levels -3 to 0 to Cluster 1 and 1 to 3 to Cluster 2. Strategy 2 groups PA levels -3 to 1 in Cluster 1 and 2 to 3 in Cluster 2.



(a) Components and eigenvalues (b) Components and explained variance

Figure 4.10: PCA results: eigenvalues and explained variance.

The dataset is tested using various typical machine learning models, employing 5-fold cross-validation to ensure the reliability of the results. The models include

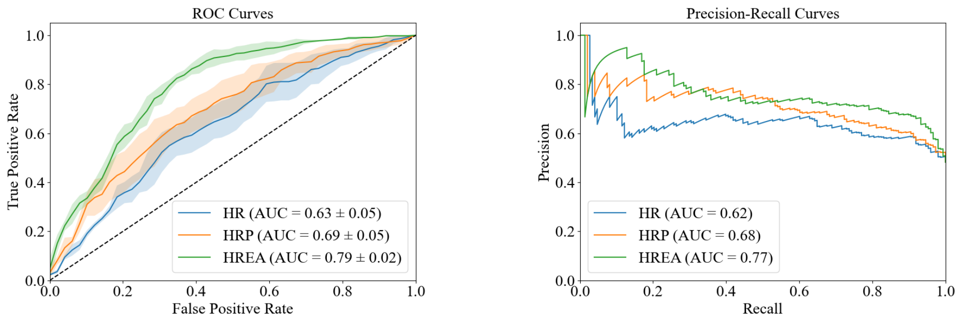
Random Forest (RF), Gradient Boosting Decision Trees (GBDT), AdaBoost with decision trees (AdaB), Feedforward Neural Network (FNN), and Extreme Gradient Boosting (XGB). The performance was evaluated based on accuracy, precision, recall, and F1 score, as shown in Table 4.3. Strategy 2 displays more balanced performance metrics between Clusters 1 and 2 across all classifiers, indicating a more uniform data separation, and is selected for further analysis.

Table 4.3: Performance metrics over strategies (Str.) and models (Mod.), with data grouped by clusters (Clus.). The metrics include accuracy (Acc.), precision (Prec.), recall (Rec.), and the F1 score (F1).

Str.	Mod.	Clus.	Acc.	Prec.	Rec.	F1
1	RF	1	0.78	0.64	0.78	0.70
		2	0.49	0.67	0.49	0.57
1	GBDT	1	0.74	0.65	0.74	0.69
		2	0.55	0.65	0.55	0.59
1	AdaB	1	0.70	0.62	0.70	0.66
		2	0.51	0.60	0.51	0.55
1	FNN	1	0.74	0.65	0.74	0.69
		2	0.55	0.65	0.55	0.59
1	XGB	1	0.78	0.64	0.78	0.70
		2	0.51	0.67	0.51	0.58
2	RF	1	0.74	0.60	0.74	0.66
		2	0.61	0.75	0.61	0.67
2	GBDT	1	0.70	0.60	0.70	0.65
		2	0.63	0.73	0.63	0.68
2	AdaB	1	0.63	0.55	0.63	0.59
		2	0.59	0.67	0.59	0.63
2	FNN	1	0.70	0.61	0.70	0.65
		2	0.65	0.73	0.65	0.68
2	XGB	1	0.70	0.58	0.70	0.63
		2	0.60	0.72	0.60	0.65

ABLATION STUDY

An ablation study is conducted to understand the impact of various feature sets on the performance of PA detection. Based on the assumption that emotional and attentional features may act as intermediate variables that improve PA detection, participants' self-reported emotions and attention are included for comparisons. Since emotional and attentional features are not typically observable in real-world conditions, we also predict these features based on human and robot movement data, achieving a valence detection accuracy of 71.0%, arousal detection accuracy of 69.2%, and attention detection accuracy of 72.8%. We evaluate the effectiveness of the following feature combinations: Robot and human motion (RH), Robot and Human motion + predicted emotion and predicted attention (RHP), Robot and human motion + human reported emotion and attention data (RHEA).



(a) ROC curves (b) Precision/recall curves

4

Figure 4.11: ROC and Precision/Recall curves showing the PA detection performance with different features.

The effectiveness of the XGBoost algorithm across a variety of feature combinations is assessed. Results include the Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves (Figure 4.11). The ROC curve (Figure 4.11a) illustrates the model's sensitivity (true positive rate) versus its fall-out (false positive rate) at various thresholds. This offers a comprehensive evaluation of the model's discriminative capability, unaffected by class distribution. The PR curve (Figure 4.11b), on the other hand, provides insight into the model's precision in relation to its recall, which is especially relevant for imbalanced datasets. The HREA feature set demonstrates the highest performance, achieving the top Area Under the Curve (AUC) scores in both the ROC (0.79 ± 0.02) and PR (0.77) analyses. This indicates a consistent and reliable detection of true positives across different thresholds, thereby enhancing the detection's reliability and robustness. The high AUC values suggest a model that accurately predicts while maintaining a balanced sensitivity and precision. In comparison, the HRP feature set shows significant improvement over the HR set, as indicated by higher AUC values in the ROC (0.69 ± 0.05 vs. 0.63 ± 0.05) and PR (0.68 vs. 0.62) curves. This improvement highlights the benefit of including predicted emotional and attentional features for superior detection capabilities.

The statistical significance of these findings is supported by the hypothesis tests detailed in Table 4.4a. Comparing HREA to HR, and HRP to HR, the analysis reveals significant t-values and small p-values, indicating not just statistical significance but also meaningful effect sizes as measured by Cohen's d. This confirms that the performance differences are statistically significant and practically important. Furthermore, the accuracy and F1 score statistics shown in Table 4.4b corroborate these results, with HREA outperforming and HRP significantly better than HR. These statistics, together with insights from the ROC and PR curves in Figure 4.11, underscore that the inclusion of emotional and attentional features, as represented by HREA, and the predicted emotional and attentional features, as described by HRP, markedly improve the model's detection effectiveness. Specifically, the high PA detection performance with the inclusion of the human self-reported emotion and attention strongly demonstrates an upper bound on the usefulness of emotion and attention to predict PA.

(a) Hypothesis tests for PA detection with different sets of features. Since the data used are normal, Cohen's d is used to measure the effect size.

Feature	Accuracy			F1 score		
	$t(4)$	p	d	$t(4)$	p	d
HR vs HREA	12.00	0.004	4.37	12.82	0.003	4.10
HR vs HRP	6.831	0.019	2.00	6.35	0.003	1.56
HRP vs HREA	8.454	0.039	1.53	8.21	0.033	1.96

(b) PA detection statistics for Accuracy and F1 score with different features.

Features	Accuracy		F1	
	μ	σ	μ	σ
HR	0.63	0.05	0.68	0.05
HRP	0.68	0.05	0.72	0.05
HREA	0.79	0.05	0.79	0.04

Table 4.4: Descriptive statistics and hypothesis tests for Accuracy and F1 score with different feature sets using the XGBoost algorithm.

4.4.3. LIMITATIONS AND DISCUSSION

There are several limitations that could potentially affect the accuracy and utility of the PARSNiP dataset. First, the reliance on participants' self-reported data for PA, emotion (valence and arousal), and attention after all interactions may introduce recall biases, potentially compromising data reliability. Although measures such as playing recorded interactions with their facial expressions and motions to the participants and maintaining brief, straightforward interactions were implemented to mitigate this issue, recall biases may still occur. Future research should consider employing real-time annotation methods like the 'think-aloud' technique, which would allow participants to report their states immediately during or just after each interaction. Second, we make the assumption of consistent human attention, emotion, and PA during 2-second intervals, which may not accurately reflect real-world dynamics. Future research should explore how varying time windows for reporting attention, emotion, and PA might more accurately represent real-life conditions. To note, the current study employs a 2-second interval to enhance the efficiency of data collection. In real-life applications, it is suggested that researchers and engineers train the system offline, subsequently applying a sliding-window approach that involves overlapping data windows that shift incrementally over time (such as per 0.5s) [328], allowing the robot to continually detect PA and adjust its behavior based on ongoing interactions. This approach aims to minimize the latency in PA detection, thereby allowing robots to adjust their navigational behavior in accordance with human preferences in time. While using an onsite camera for data collection is advantageous for controlling experimental conditions, we recognize that the specific setup—positioning the camera on a table at 0.75m height and 1m distance—may not fully align with real-world applications. To bridge this gap, several strategies can be employed to adapt the collected data for use with different setups such as onboard cameras. For example, data mapping and transformation techniques can simulate the perspective of an onboard camera, while data augmentation and simulation can create additional datasets that replicate various camera positions and angles. These approaches can improve the practicality of the dataset and deserve further investigation.

4.5. Conclusion

We introduced a novel dataset, i.e., PARSNiP, which offers a complete range of PA and incorporates emotional and attentional features to enhance PA detection. Firstly, 30 participants interacted with a mobile robot 8 times, where an onsite 3D camera and robot odometry were used to record human and robot motion data. Then, these data were processed to compute motion features, while the recordings showing the participants' front views were replayed to the participants to enable recall and report of their PA, emotion, and attention to enable dataset creation. Finally, the dataset was tested with several typical machine learning models, and an ablation study was conducted to compare PA detection performance among different input feature groups: motion features alone, motion features with predicted emotion and attention, and motion features with the participants' self-reported emotion and attention. The results showed improved PA detection when including emotion and attention, even if they are predicted. This underscores the critical role of human emotion and attention in enhancing the accuracy of PA detection. This dataset can be used to train machine learning models, which can then be implemented in robots to adapt their behaviors—such as stopping or moving away that are typical robot adaptations— when around humans, enhancing their alignment with human expectations and social acceptance. There are several limitations, including the assumption of constant human attention, emotion, and PA during 2-second intervals, potential biases in participants' retrospective reports of their states, and direct applicability on the robot. Future research should investigate ways to collect real-time data and employ onboard cameras to simulate real-world conditions better and improve PA detection accuracy.

5

ADAPTING INAPPROPRIATE ROBOT NAVIGATION BEHAVIOR

When navigating social environments, robots inevitably exhibit behaviors that are perceived as inappropriate by humans. Current robots lack the ability to adapt to inappropriate behaviors, resulting in their repeated and continued occurrence. Therefore, it is important to answer the research question: How do humans prefer a robot to adapt its inappropriate navigation behavior in narrow environments (RQ4)? This study explores human-preferred adaptations to inappropriate robot navigation behaviors through a series of human-robot interactions, a semi-structured interview, and an online survey. 12 participants were recruited to interact with a mobile robot at virtual pathway intersections in a lab, reporting 139 instances of inappropriate robot behaviors. A subsequent semi-structured interview regarding these instances yielded 9 types of inappropriate robot behaviors and 10 major types of human-preferred robot adaptations, ranging from general ones, such as stopping the motion, to more specific ones, like moving away and then stopping. One week later, an online survey was conducted, presenting 12 selected human-preferred adaptations to the same participants to evaluate how effectively these addressed the inappropriate behaviors they had previously identified. The results highlight a strong human preference for the robot to move to the side of the pathway and stop in most scenarios. This might serve as a general adaptation for addressing inappropriate robot navigation behaviors, especially in narrow environments.

5.1. INTRODUCTION

Robots play a vital role in social settings, ranging from providing services and assisting in disaster recovery to transforming healthcare delivery [335]. Central to these operations are socially aware navigation approaches, which aim to enable

This chapter has been published as: Zhou Y, Vroon J, Kortuem G. Exploring Human Preferences for Adapting Inappropriate Robot Navigation Behaviors: A Mixed-Methods Study[J]. IEEE Robotics and Automation Letters, 2024.

robots to navigate in a safe, effective, and socially acceptable manner [11]. Key to socially aware navigation is the adherence to established social norms, which includes maintaining safe interpersonal distances [39], minimizing disruptions within crowded spaces [8], and giving precedence to human activities [336]. Moreover, these robots integrate advanced path prediction and planning techniques to improve their interaction with humans [97, 337]. Their ability to navigate socially is further enhanced by recognizing and adapting to the nuanced psychological states of humans, including emotional states [165], dominance levels [52], and politeness [19]. These developments have increasingly enabled robots to navigate socially and safely around humans.

While the majority of these approaches have dominantly focused on perfecting robot behaviors, some other studies have tried to resolve challenges that emerge during navigation when applying these approaches. Conflicts have been a significant challenge in socially aware navigation. According to Mirsky et al.: "A conflict between a robot and [...] pedestrians is a situation in which if there is no change of direction or speed by at least one of the parties, they will collide" [338]. One common strategy for resolving conflicts is for the robot to adjust its path or speed based on the inferred intentions or movements of the humans [339]. Other studies focused on enhancing the robot's communication capabilities, using direct communications such as visual signals (e.g., lights or projections) or sounds, thus preventing misunderstandings and potential conflicts [340]. Some studies have tried to address the "freezing robot problem (FRP)" [341], which occurs when a robot navigating a crowded environment becomes paralyzed or stuck. One method, Frozone, computed a deviation velocity that avoids the Potential Freezing Zone (PFZ) to ensure smooth and collision-free navigation [59]. Additionally, a recent study tackled the FRP problem by utilizing deep reinforcement learning, incorporating spatial-temporal reasoning along with real-time pedestrian speed data [60]. Narrow spaces present a distinct problem for robots as they offer limited space for both the robot and human to move around with ease. Senft et al. found that humans generally favor robots that rotate their bodies to clear the path, as opposed to a sliding motion [298]. In an effort to understand the impact of inappropriate robot behaviors on humans, Koay et al. conducted human-robot interaction studies and identified a set of potentially discomforting robot behaviors, such as getting too close or blocking paths [323]. However, all these works did not investigate human perception of robot behaviors, which is crucial for pinpointing truly inappropriate behaviors. A recent study simulated robot actions and evaluated their social appropriateness using human assessments. The trained model effectively predicted social appropriateness, but the simulated, task-oriented nature of these actions limits real-world relevance [342]. Another study by Vroon et al. introduced the concept of "perceived appropriateness" (PA), which refers to an individual's subjective perception of how appropriate a robot's behavior appears [10]. They conducted real-life experiments to create a dataset for detecting human PA of robot positioning behavior but ignored other inappropriate behaviors like path blockage and acceleration and ignored robot adaptations.

Though not in the navigational contexts, many studies in human-robot interactions have widely investigated inappropriate robot behaviors (also known as "errors", "mistakes", and "failures") [61, 343]. The researchers identified various types of inappropriate behaviors in robots, which were primarily categorized into two broad groups: technical failures (comprising software and hardware issues) and social norm violations. Humans were found to exhibit varied social signals when confronted with different types of robot errors [319], which enabled the robot's detection of their errors [68]. For instance, during technical failures, individuals showed significantly more facial expressions, head movements, body movements, and gaze shifts compared to during social norm violations [64]. These inappropriate robot behaviors were found to decrease the perceived trustworthiness and reliability of robots [65]. Consequently, numerous studies have explored robot adaptations that could mitigate such negative perceptions, revealing human preferences for error-specific adaptations [63]. While the studies were carried out in non-navigational settings, potentially limiting their direct relevance to socially aware navigation, their insights into human preferences for error-specific robot adaptations might still hold significant value in navigational contexts.

In conclusion, existing studies have 2 gaps in the socially aware navigation context that this study seeks to bridge. The first gap is the limited understanding of human perception of robot navigation behaviors. We propose to address the research question: **RQ4.1: "How do humans perceive the appropriateness of robot navigation behaviors, how do they trigger human concerns?"** This question further breaks down into specific sub-questions: **RQ4.1.1: "What factors contribute to the emergence of perceived inappropriate robot behaviors?" RQ4.1.2: "How do humans perceive the appropriateness of robot navigation behaviors?" RQ4.1.3: "What are the human concerns with the perceived inappropriate robot navigation behaviors?"**

The second gap is a lack of knowledge regarding how humans prefer robots to adapt their inappropriate navigation behaviors. As the conflicts underlying inappropriate behaviors are much more frequent in narrow environments [80], we propose to address the research question: **RQ4.2: "How should robots adapt to resolve inappropriate navigation behaviors in narrow environments?"** This question further breaks down into specific sub-questions: **RQ4.2.1: "How do humans prefer the robot to adapt its inappropriate navigation in narrow environments?" RQ4.2.2: "How do humans evaluate the appropriateness of robot behavior after it behaved inappropriately in narrow environments?" RQ4.2.3: "How different adaptations are preferred under different types of inappropriate behaviors in narrow environments?"**

5.2. METHOD

5.2.1. SAMPLING AND RECRUITMENT

A purposive sampling strategy was employed for participant recruitment. 12 participants (mean age = 28.8 years, SD = 6.0 years; 8 females and 4 males) were recruited for the experiment. Each participant received a €10 voucher upon completing the entire experiment. Participants were selected based on their experience in relevant fields of design, particularly those involving interaction design and human-centered design, as these areas provide essential expertise in interactions and problem-solving, which are crucial for exploring adaptations to inappropriate robot behaviors. Eligibility also requires physical capability to interact with the robot and proficiency in English for clear communication. Participants were neither associated with the researchers' laboratory nor briefed on the specific research hypotheses before. Recruitment of additional participants was stopped when thematic saturation was reached, indicating that no new themes emerged during the analysis [344]. The research protocol was approved by the university's Human Research Ethics Committee.

5.2.2. Experimental procedure

The experimental procedure is presented in Table 5.1 and consists of 3 steps: human-robot interactions, a semi-structured interview, and an online survey.

Table 5.1: Experimental procedure: the purpose and data collection for each step.

Step	Purpose	Data collection	RQ
Human-robot interactions	Report inappropriate robot behavior	"Press the button in your hand if you perceive that the robot behaves inappropriately". Multiple button presses in close succession are treated as a single instance.	
Semi-structured interview	Robot behavior factors	Q1.1: What occurred before you pressed the button?	4.1.1
	PA type	Q1.2: Which aspect of the robot's behavior was inappropriate?	4.1.2
	Concern	Q1.3: Why did the robot's inappropriate behavior concern you?	4.1.3
	Preferred adaptation	Q2.1: After pressing the button, how would you have preferred the interaction to continue?	4.2.1
	Evaluation of the robot's next action	Q2.2.a: What did the robot do after the button press? Q2.2.b: (-2 to 2, 5-point scale): How appropriate did you find the robot's behavior after the button press?	4.2.2
Online survey	Evaluation of 12 adaptations	Q2.3 (-2 to 2, 5-point scale): For each adaptation, how appropriate would you find it in this situation?	4.2.3

Step 1. Human-robot interactions

Upon arrival, the participants were briefed on the study, consent was obtained, and they were introduced to the experiment: They would walk through the pathway intersections in an indoor lab with designated start and end points 6 times (see Fig. 5.1), each time interacting with a Clearpath Jackal robot (0.51 m x 0.43 m x 0.25 m); Due to the nature of the crossings, both the robot and the participants were still allowed to move out of the boundary whenever they felt the need; During interactions, they would press a button in their hands whenever they perceived the robot behaving inappropriately and needed to make adaptations [10].

These 6 interactions altogether exposed participants to a total of at least 21 distinct robot behaviors, detailed in Table 5.2, where the protocol for the operator and

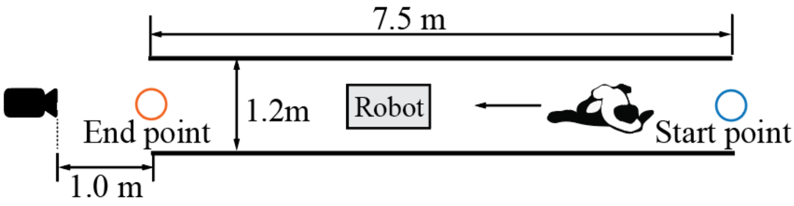


Figure 5.1: Overview of the designed setting.

tasks for the participants were also explained during each interaction. The operator controlled the robot to achieve the desired robot behaviors without any feedback information from the participants' button presses. Should there be any danger, the operator will initiate the emergency stop. These behaviors were inspired by rich literature discussed in Section 5.1, containing identified inappropriate behaviors from the literature and designed inappropriate behaviors by deliberately violating social norms widely applied in socially aware navigation approaches. The main pathway, with a length of 4.8 m and a width of 1.2 m, was designated for interactions 1-5, covering behaviors B1-B17. The entire region (including the main pathway), with pathways of varying widths (0.80 m, 1.00 m, and 1.20 m), was specifically used for interaction 6, covering behaviors B18-B21.

5

The robot was controlled by a skilled operator via a Bluetooth PS4 controller, employing the Wizard of Oz protocol [326]. The operator's station, equipped with a standing desk, a chair, and a laptop, was located in a corner of the room. The setup was designed so that the participant could not see the operator, thus ensuring adherence to the Wizard of Oz protocol. The interactions were recorded by a camera positioned facing the main pathway. The teleoperator was trained across all 6 interaction types. For "Follow", the operator practiced speed and acceleration control and maneuvers for human overtaking and catching up. In "Avoid", training focused on alternating between moving aside and retreating. "Get close" involved personal space violations with sudden and slow stops. "Sudden motion" emphasized path narrowing, moving towards humans, and sudden direction changes. "Overtake" centered on speed control and safe overtaking, while "Random motion" included varying speed, direction, and extended stops over 3 minutes.

STEP 2. SEMI-STRUCTURED INTERVIEW

After 6 interactions in Step 1, the participants were presented with the recorded videos (with timestamp information) of the interactions and the timestamp of their own button presses to pinpoint and recall each instance. They were allowed to play the videos before and after the moment of their button press to help understand the situation better. After identifying each button press, they were interviewed concerning the robot behavior factors (Q4.1.1), the inappropriate robot behavior (Q4.1.2), their underlying concern (Q4.1.3), preferred robot adaptation (Q4.2.1), and the evaluation of the robot's next action (Q4.2.2). After the interview, the participants were also asked how they would agree if the robot could adopt a general adaptive

Table 5.2: Summary of 6 interactions, robot behaviors, and the protocol for the operation

Interaction	Behavior	The protocol for the operator and tasks for the participants. RRobot and human points are visualized in Fig 5.1. RA, RB, RC, and RD refer to the robot points, while HA, HB, HC, and HD refer to human points. For each interaction, both the robot and the human start moving simultaneously.
1: Follow	B1: Follows slowly [345] B2: Follows at a similar speed [345] B3: Follows at a faster speed [338] B4: Blocks from the front [323] B5: Accelerates $\leq 3\,\mathrm{m/s^2}$ [31, 346] B6: Moves slowly [153] B7: Is overtaken [347]	The robot travels RA to RB and back, the human HA to HB and back, each completing one round trip in the main pathway. The robot moves slowly towards point RB to allow the human to overtake it (B1, B4, B6, B7), then accelerates to catch up (B2, B3, B5). At point RB, the robot turns and repeats this sequence: waiting for the human to pass (B1, B4, B6, B7) and then accelerating to catch up (B2, B3, B5).
2: Avoid	B8: Avoids from the side [323] B9: Avoids by retreating [323]	Both the robot and human make 2 round trips between RA and RB and HA and HB respectively in the main pathway, with 4 encounters. The robot alternates moving aside and retreating (B8 & B9) for each encounter.
3: Get close	B10: Violates personal space [323] B11: Suddenly stops when close [323] B12: Slowly stops when close [323]	Both the robot and human make 2 round trips between RA and RB and HA and HB, respectively in the main pathway, with 4 encounters. The robot alternates, stopping quickly (B11) and slowly (B12) when near the human (B10).
4: Sudden motion	B13: Narrows human path [347] B14: Moves toward human [31] B15: Suddenly changes direction [153]	Both the robot and human make 2 round trips between RA and RB and HA and HB, respectively in the main pathway, with 4 encounters. The robot narrows the path (B13), approaches the participant (B14), and then suddenly changes the direction (B15).
5: Overtake	B16: Moves fast at $1-2\,\mathrm{m/s}$ [338, 341] B17: Overtakes the human [347]	The robot travels RA to RB and back, the human HA to HB and back in the main pathway, each completing one round trip. The robot alternates moving fast and overtaking the human (B16, B17) for each encounter.
6: Random motion	B18: Moves towards the human [323] B19: Blocks from the side [323] B20: Changes direction [31] B21: Stops for a long time [341]	The robot and human each navigate 4 points until the 3-minute timer alarms. The robot follows the sequence RA to RD and back repeatedly, while the human follows the same pattern from HA to HD. The robot randomly changes speeds and directions (B18, B19, B20) and stops for a long time (B21).

motion that worked for most of the inappropriate behaviors they experienced. They were also asked whether and what external interface(s) might be needed along with the robot's adaptive motion to improve their perception of the robot.

After all interactions and the interview, the interview answers to Q4.2.1 concerning human-preferred adaptations were coded and analyzed. Among the 34 preferred adaptations (see the Appendix 5.6.1), 12 were selected based on participants' answers to proceed to Step 3, including 11 most frequently reported adaptations and 1 baseline adaptation of "continue".

STEP 3. ONLINE SURVEY

One week after completing all interactions and interviews, the same 12 participants were invited for a follow-up online survey. The online survey included recordings of the participants' interactions with the robot and the timestamps of their button presses, enabling them to pinpoint and recall each instance of their own button press. They were asked to imagine that each of the 12 adaptations would occur immediately after they pressed the button. Each adaptation was provided with explanations to assist the participants in understanding the adaptation. They then rated how appropriate each adaptation would be for each inappropriate robot behavior previously identified by their own (Q4.2.3, 5-point scale, from strongly inappropriate to strongly appropriate).

5

5.2.3. DATA CODING AND THEME EXTRACTION

The interview data were transcribed verbatim into Microsoft Word and systematically analyzed using ATLAS.ti, following Braun & Clarke's six-phase inductive approach to thematic analysis [88]. Initially, the primary researcher, Yunzhong Zhou (Y.Z.), became familiar with the data by reading the transcripts multiple times and noting initial ideas and reflections. In the second phase, Y.Z. generated initial codes by identifying meaningful segments throughout the data. To enhance coding reliability, 10% of the randomly selected data was coded by an independent researcher in human-robot interaction studies. Discrepancies were discussed and resolved, which further refined the coding framework. Cohen's kappa values ranged from 0.71 to 1.00, indicating substantial to perfect agreement [348]. Third, Y.Z. grouped related codes into potential themes. These themes were then collaboratively reviewed and refined with the second author, Jered Vroon (J.V.), resulting in a preliminary thematic map. The fourth phase involved a comprehensive review of these themes, assessing them against the coded extracts and the entire dataset. This review led to the modification of some themes, including merging related themes and redefining others. Themes were defined and named through iterative discussions involving Y.Z., J.V., and another author, Gerd Kortuem (G.K). Finally, a scholarly report was produced.

The objective robot-human relative poses (including position and orientation [212]) in the recorded videos were also coded, and each encounter was associated with a timestamp. As the precise timing of participants pressing the button varied (i.e., some pressed it after the encounter, some during), a combination of the video and the interview was used to pick the precise moment to code (within 2 seconds of the timestamp, at exactly the timestamp if unclear). Specifically, the robot's relative poses to the participants were coded, including relative position (front, behind, left, right) and orientation (same, opposite, left, right).

The coding and the results from the interview and videos are presented in Table 5.3. Through the interview, participants altogether reported 139 instances of inappropriate robot behaviors. Typically, participants provided 1 answer per instance for most questions, except for their concerns in Q4.1.3 and the preferred adaptations in Q4.2.1. This resulted in 139 instances of inappropriate robot behaviors and evaluations of the robot's next action, alongside 148 instances of preferred adaptations and 203 instances of concerns.

Table 5.3: Coding data overview. K denotes Cohen's kappa value [348], and Num denotes the number.

Purpose	Source	Type	Num	K
Robot behavior factors	Videos, Interview	Objective	139	0.71
	Interview (Q4.1.1)	Subjective	139	0.73
Inappropriate behavior	Interview (Q4.1.2)	Subjective	139	0.86
Concern	Interview (Q4.1.3)	Subjective	203	0.71
Preferred adaptation	Interview (Q4.2.1)	Subjective	148	0.79
Robot's next action	Interview (Q4.2.2a)	Subjective	139	0.86
Evaluation of the robot's next action	Interview (Q4.2.2b)	Subjective	139	1.00
Evaluation of 12 adaptations	Online survey (Q4.2.3)	Subjective	139	1.00

5.3. Results

5.3.1. Inappropriate robot behaviors

Robot behavior factors

Robot behavior factors contributing to inappropriate robot behaviors include both objective robot-human relative poses and subjective reports from the participants.

An agent's pose comprises position and orientation [212]. In our coding, the robot-human relative poses consist of 16 types, including combinations of different positions (front, behind, left, right) and orientations (same, opposite, left, right). As can be seen from Fig. 5.2, the relative poses strongly influence the number of reported inappropriate robot behaviors, especially when the robot approaches the participants, particularly from the front.

Furthermore, without being directly asked in Q4.1.1, the participants reported 3 key robot behavior factors that influenced their button presses: acceleration (43 instances), smoothness (42 instances), and speed (14 instances). Regarding acceleration, feedback primarily focused on the robot accelerating (37/43) rather than decelerating, especially when the robot approached or followed them from the front (35/37). The smoothness factor, described by the participants as "random", "unpredictable", or "jerky", was majorly mentioned when the robot was either in front of or behind the participants (35/42). The speed factor was often associated

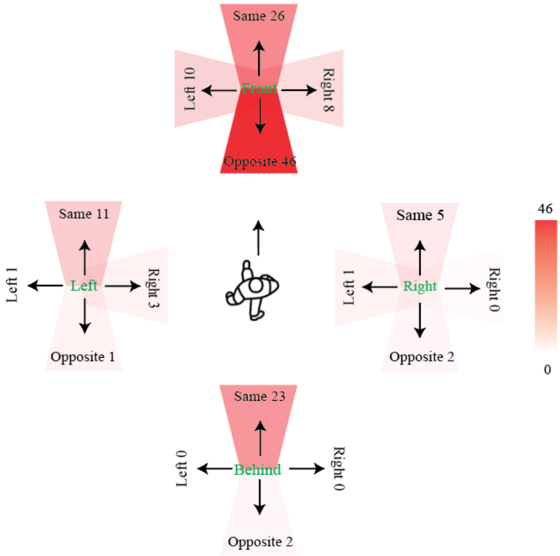


Figure 5.2: Number of the reported inappropriate robot behaviors across different robot-human relative poses.

with the robot moving too slowly or remaining stationary (12/14) rather than moving too fast.

INAPPROPRIATE BEHAVIOR TYPES

Fig. 5.3 shows 139 instances of inappropriate robot behaviors reported by 12 participants, categorized into 9 types. The number of instances reported by each participant ranges from 2 (P11) to 18 (P4 and P8). While most participants reported multiple types of inappropriate behaviors, P11 was only sensitive to “Accelerate", and P4 and P9 were dominantly sensitive to “Block". Among the various types of inappropriate behaviors, “Block," where the robot inappropriately obstructed the participants' path, was the most frequently reported (59/139). This happened predominantly when the robot was positioned in front of the participants (48/59), especially in parallel (same or opposite) orientation (36/48), as shown in Fig. 5.4. In such scenarios, the participants often perceived the robot as constantly blocking (19/59) or making unexpected blocks (12/59). “Accelerate" was also frequently reported (48/139), especially when the robot was in front (20/48) or behind the participants (17/48), see Fig. 5.4. In such scenarios, the participants often reported that the robot was making unexpected accelerations (21/48). Others include the robot moving “Towards" the participants, exhibiting “Unpredictable" or “Sudden" movements, “Follow", getting “Too close", “Watch", or “Squeeze" the participants' path. Specifically, “Watch" denotes the participants' feeling of being watched or observed, while “Squeeze" refers to the robot restricting the participants' paths.

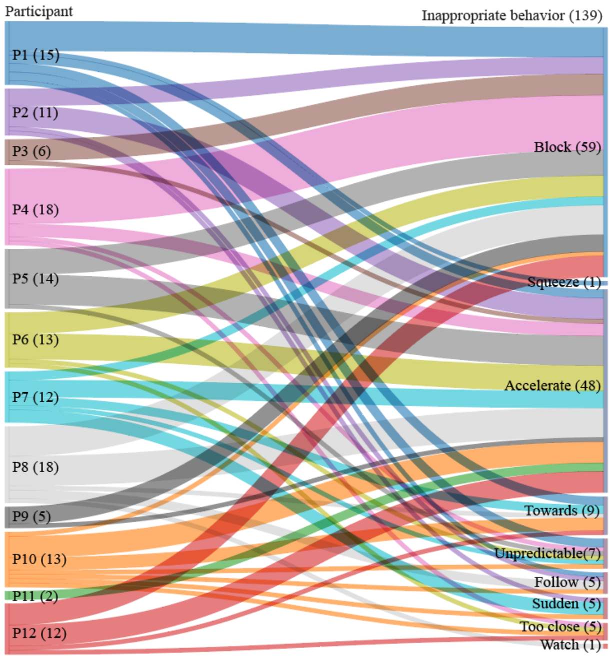


Figure 5.3: Reported inappropriate robot behaviors (with the number).

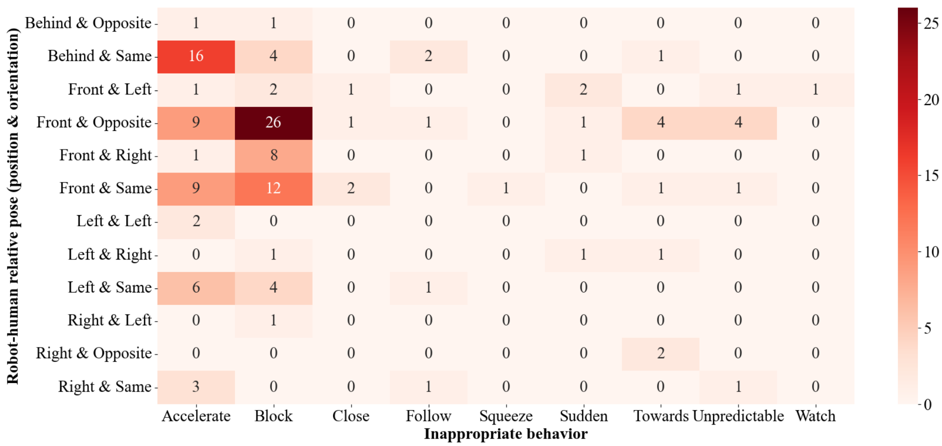


Figure 5.4: Heatmap of inappropriate behavior types vs. Robot-to-human pose

CONCERNS

Table 5.4 presents 203 instances of concerns from the participants, categorized into 8 types. The most common concern was “Safety" (79 instances), which was mostly related to the robot being in a collision course with the participants (relative poses: front & opposite (25 instances), behind & same (16 instances), and front & same (12 instances)). The concern of the robot “Interfere" with their walking paths (38 instances) was also frequently reported, especially when the robot was in the front (relative poses: front & opposite (15 instances), front & same (9 instances), and front & right (4 instances)). “Confuse" concerns (33 instances) mostly arose when the robot was in the front (relative poses: front & opposite (12 instances), and front &same (9 instances)). Other concerns include “Disrespect", “Comfort", “Efficiency", a violation of “Privacy", and increased “Cognitive load".

Table 5.4: Number of Concerns

Concern	Number
Safety	79
Interfere	38
Confuse	33
Disrespect	25
Comfort	13
Efficiency	5
Privacy	5
Cognitive load	5

Table 5.5: Number of Preferred Adaptations

Preferred adaptation	Number
Away	60
Stop	37
Slow	19
Away, stop	11
Slow, speak	9
Stop, speak	7
Predictable	3
Away, speak	1
Continue	1
Stop, light	1

5

5.3.2. Adapting inappropriate robot navigation

Preferred adaptations

Table 5.5 shows 148 instances of preferred adaptations categorized into 10 major types (see Appendix 5.6.1 for all 34 types of preferred adaptations). The most common was for the robot to move “Away" (59 instances), especially when the robot was in the front, especially with the relative orientation of opposite (25 instances) or same (12 instances). The second most frequently reported was to “Stop" (37 instances), especially when the robot was on a collision course with the human (relative poses: front & opposite (8 instances), behind & same (7 instances)). Others include for the robot to “Slow down", “Away, stop", “Slow, speak", “Stop, speak", make “Predictable" behavior, “Away, speak", “Continue", or “Stop, light".

The relationships between inappropriate behavior, concern, and preferred robot adaptation have also been analyzed and presented as a Sankey diagram in Fig. 5.5. Different colors in the diagram represent different types of inappropriate robot behaviors. For analytical clarity and focus, only those sub-categories with at least 5 instances were included, collectively representing over 95% of the cases within each category. Additionally, to facilitate a clear understanding of human preferences in response to inappropriate behaviors, the preferred adaptations were clustered into 3 categories: move “Away", “Slow down", and move “Away and slow down".

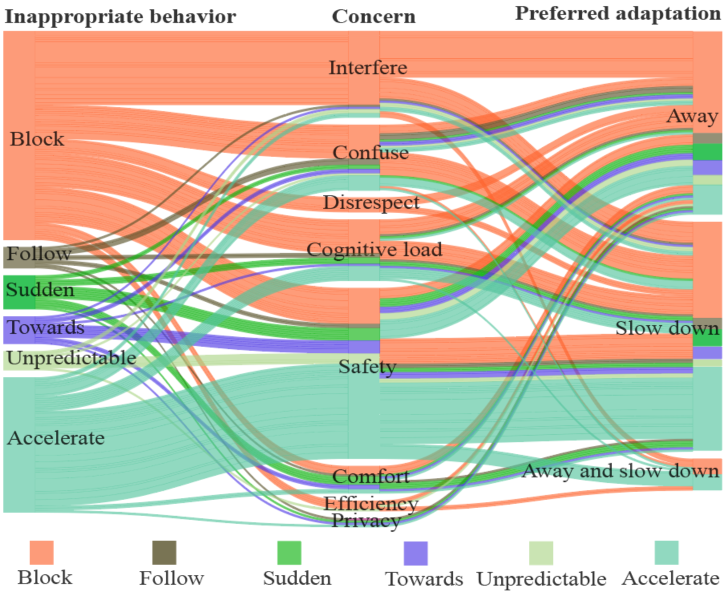


Figure 5.5: The Sanky diagram showing relations between inappropriate behavior, concern, and preferred adaptation. Different colors represent different inappropriate behavior types.

From the inappropriate behaviors perspective, “Block" and “Accelerate" predominantly trigger a broad range and a significant proportion of human concerns. Notably, “Block" triggered concerns that were more evenly distributed across “Interfere", “Confuse", “Disrespect", and “Safety", while “Accelerate" primarily triggered “Safety" concerns. Other types were associated with different concerns such as “Sudden" behavior causing concerns of “Safety", “Confused", or “Comfort".

From the concern perspective, the same concern could be triggered by different inappropriate behaviors. Notably, “Interfere" was dominantly triggered by “Block", and “Safety" was largely triggered by “Accelerate". These different concerns also influenced humans’ preferred robot adaptations, as evidenced by a strong human preference for the robot to “Slow down" rather than to “Move away" when having “Safety" concerns and the preference for the robot to move “Away" rather than to “Slow down" when having “Interfere" concerns. For other concerns, participants showed a relatively balanced preference for the robot to either move “Away" or “Slow down". Notably, the preference of the robot to move “Away" addressed a wider range of concerns than to “Slow down" (8 vs. 6 types).

The direct relationship between inappropriate behaviors and preferred adaptations was further analyzed, which can be seen from the distribution of the color blocks in columns of the preferred adaptations category in Fig. 5.5. Most inappropriate behavior types led to an evenly distributed preference for adaptations of “Away" and

“Slow down." However, one notable exception was the “Accelerate", with the majority leading to the preference for the robot to “Slow down" rather than to move “Away".

Evaluations of the Robot’s Next Action

The participants also reviewed and evaluated the appropriateness of the robot’s next action. Interestingly, participants often perceived the robot as adapting its behavior (96 instances), although the operator did not receive any information from them and thus did not adapt accordingly. This included “Slow down" (27/139), “Turn" (27/139), “Stop" (13/139), “Jerky" (13/139), “Accelerate" (12/139), and “Follow" (4/139). In only 43 instances was the robot perceived as “Continue" its previous behavior.

The mean ratings for the robot’s next actions are presented in Table 5.6. The robot’s next actions were generally rated negatively by the participants, with a mean score of -0.47 and an std of 0.98. Actions of “Continue" were largely rated negatively across various types of inappropriate behaviors, with a mean score of -0.91 and a std of 0.32. Other actions, such as “Slow down" or “Stop" received mixed evaluations depending on the situation. Notably, “Slow down" following the “Block" was slightly rated positively (mean: 0.43, std: 0.51), whereas “Slow down" following “Follow" was evaluated negatively (mean: -0.33, std: 0.47). Meanwhile, “Stop" following “Block" or “Towards" motion was generally evaluated positively, while “Stop" following “Accelerate" was rated slightly negatively.

5

Evaluations of 12 Robot Adaptations

Average participant evaluations of 12 adaptations concerning their own identified inappropriate robot behaviors through the online survey are shown in Table 5.7. The adaptations include: “Continue”, where the robot continues its current motion; “Away”, where the robot moves away from the participants; “Slowly away”, in which the robot moves away at a slow pace; “Retreat”, where the robot backs off; “Retreat, stop”, where the robot backs off and then stops; “Aside”, where the robot moves to one side of the pathway; “Aside, stop”, where the robot moves to one side of the pathway and then stops; “Stop”, where the robot stops; “Quickly stop”, where the robot stops quickly; “Slow”, where the robot slows down without stopping; “Slow, inform”, where the robot slows down and informs its next action; and “Predictable”, where the robot moves in a predictable manner, i.e., in a straight line.

In most scenarios, the robot continuing its previous inappropriate behaviors consistently received negative evaluations (mean -1.43, std 0.42). Although the robot received a positive evaluation during the “Squeeze" situation, this was based on a single instance and might represent an outlier, as the robot generally received negative evaluations overall. Adaptations such as “Slowly away" and moving “Aside" or “Aside, stop" generally received positive evaluations across all situations, which might indicate human preferences for these adaptations. Among the 12 adaptations, “Aside, stop" was mostly favored, receiving the highest average evaluations across most inappropriate behavior types (mean 1.19, std 0.35). A Mann-Whitney U test was further conducted comparing “Aside, stop" with other adaptations due to the non-normal distribution of the data [333]. The results revealed significant differences between “Aside, stop" and all other adaptations, with Bonferroni-corrected p-values

Table 5.6: Average evaluation of the robot's next action following the PA, along with the number of instances.

PA type (#)	The robot's next action (#)						
	Continue (43)	Slow down (27)	Turn (27)	Stop (13)	Jerky (13)	Accelerate (12)	Follow (4)
Block (59)	-0.86 (22)	0.40 (5)	-0.12 (16)	0.50 (4)	-0.75 (8)	-1.00 (3)	0.00 (1)
Accelerate (48)	-1.29 (7)	0.00 (19)	-0.20 (5)	-0.17 (6)	-0.67 (3)	-1.17 (6)	-1.00 (2)
Towards (9)	-1.33 (3)		0.00 (2)	0.33 (3)		-2.00 (1)	
Unpredictable (7)	-1.00 (4)	0.00 (1)	-2.00 (1)				1.00 (1)
Sudden (5)	1.00 (1)		0.00 (1)		0.00 (1)	1.00 (1)	1.00 (1)
Follow (5)	-1.00 (2)	-1.00 (2)	1.00 (1)				
Too close (4)	-1.00 (3)		0.00 (1)		-1.00 (1)		
Watch (1)						-2.00 (1)	
Squeeze (1)	-1.00 (1)						
Overall (139)	-0.47 (139, std: 0.98)						

Values range from strongly inappropriate (-2, deep red) to strongly appropriate (+2, deep blue).

Table 5.7: Average evaluations of 12 adaptations for each PA type, calculated based on the number of instances of that PA type.

PA type (#)	Adaptation (#)											
	Continue	Away	Slowly away	Retreat	Retreat, stop	Aside	Aside, stop	Stop	Quickly stop	Slow	Slow, inform	Predictable
Block (59)	-1.31	1.10	0.92	0.34	0.42	1.08	1.17	0.03	-0.15	-0.12	0.12	0.86
Accelerate (48)	-1.66	0.60	1.00	0.51	0.96	0.98	1.30	0.21	-0.11	0.04	0.21	0.62
Towards (9)	-1.48	0.91	0.99	0.37	0.55	1.02	1.21	0.08	-0.10	0.00	0.18	0.73
Unpredictable (7)	-1.29	0.86	0.86	0.14	0.14	1.00	1.14	-0.40	-0.40	0.00	0.00	0.60
Sudden (5)	-1.40	0.80	1.20	0.40	0.40	0.80	1.00	-0.40	-0.40	0.00	0.00	0.60
Follow (5)	-1.42	0.76	0.88	0.30	0.43	0.97	1.14	-0.20	-0.34	-0.09	0.08	0.64
Too close (4)	-1.25	0.75	1.00	0.25	0.25	1.50	1.25	0.00	0.00	0.00	0.25	0.50
Watch (1)	-1.00	1.00	2.00	1.00	1.00	2.00	1.00	0.00	0.00	1.00	0.00	1.00
Squeeze (1)	1.00	1.00	1.00	-1.00	-1.00	1.00	1.00	-1.00	-1.00	0.00	0.00	1.00
Overall (139)	-1.43	0.85	0.93	0.35	0.54	1.01	1.19	0.06	-0.16	-0.03	0.16	0.76

Values range from strongly inappropriate (-2, deep red) to strongly appropriate (+2, deep blue).

well below the adjusted significance level of $= 4.5 \times 10^{-3}$ [349]. This indicates the dominant human preference for the robot to move “Aside, stop" to adapt its inappropriate behavior.

5.4. DISCUSSION

5.4.1. INAPPROPRIATE ROBOT BEHAVIORS: FACTORS, TYPES, AND CONCERNS

This study identifies 9 types of perceived inappropriate robot behaviors (Fig. 5.3), which extends beyond the inappropriate robot positioning [10]. These inappropriate behaviors were largely affected by robot-human relative poses (as shown in Fig. 5.2), especially when a robot was in a collision course with the participants. This aligns with studies in the literature on solving potential collisions/conflicts with humans [338, 350]. Robots moving parallel to humans have led to a significantly higher number of reports of inappropriate behavior compared to those moving in other directions (95 vs. 4). This highlights the critical need to avoid parallel collision paths, an issue often neglected in current research on human-robot conflict resolution, which typically does not account for directional differences in conflicts. Furthermore, participants' perceptions of the robot were also influenced by other factors, such as motion smoothness, acceleration, and speed. Yet, due to its exploratory nature, this study did not implement stringent controls over parameters such as speed and acceleration rates. This limitation prevented the study from making in-depth quantitative insights and comparisons concerning how these factors contribute to humans' perceptions. Future research should conduct more controlled interactions with more fine-grained robot motion parameters, such as speed and acceleration rate, to allow for rigorous quantitative analyses. This will enhance our understanding of the complex dynamics in socially aware navigation, thus providing deeper insights into robot behavior design.

Among the inappropriate behaviors, “Block" and “Accelerate" were reported with far greater frequency than others and thus deserve more attention. Specifically, the inappropriate “Block" often stemmed from its recurrence or suddenness, while current literature addressing conflicts and blockage often focused on avoiding them [338, 350]. In situations where conflicts are inevitable, particularly in crowded environments, robots should balance the timing and frequency of these conflicts across different individuals. This approach helps reduce the impact on any individual and minimizes potential disruptions. Inappropriate accelerations were most frequently reported when the robot approached the participants from behind and moved towards them (see Fig. 5.4), as well as when it was positioned directly in front, which highlights the need for considering robot-human relative poses when it accelerates. Participants frequently reported that the robot's acceleration was too abrupt, a finding consistent with prior studies imposing constraints on robot acceleration rates [26, 351]. However, the specific parameters used in those studies lack sufficient empirical support. Future research could build upon these findings to explore how humans perceive varying acceleration rates, contributing to the design of more suitable robot behaviors. Moreover, some less-studied inappropriate robot behaviors were identified, such as “watch", “follow", and “squeeze", which deserve further investigation. While these behaviors were less frequently reported in this study, they might actually occur more often in other contexts. In crowded areas such as shopping malls, hotels, or public stations, a robot “watching" people could trigger privacy concerns [352], while inappropriate “squeezing" in busy streets may lead to feelings of physical threat or unease. Similarly, in museums or theme parks, service robots might “follow" visitors to offer assistance or information [353, 354], but doing so too closely or persistently could trigger discomfort or anxiety. Further studies could focus on settings where these inappropriate behaviors are more common and examine the factors that trigger them, helping design robots to avoid them. To note, participants only reported inappropriate robot behaviors after all interactions, using timestamp information to synchronize their button presses with the video recordings. This helped mitigate recall biases, but the external point of view and the delay of the reports may still cause bias issues. Future studies could explore real-time data collection methods, such as a “think aloud" approach [355], to minimize such biases. Furthermore, while this study employs manual operations to achieve the desired interaction efficiently and explore the richness of interactions, this approach introduces limitations of potential inconsistencies and reduced repeatability. Future research could build on these findings by programming the robot to standardize behaviors across participants, enabling more meaningful comparisons. To note, individual differences make achieving 100% repeatability in human-robot interactions impossible.

The inappropriate behaviors observed in this study were triggered through participants' interactions with at least 21 robot behavior types (as shown in Table 5.2). These behaviors, especially the overtaking/overtaken and sudden motion changes, enrich the understanding of potentially inappropriate robot navigation behaviors identified by Koay et al. [323]. These behaviors were integrated into 6 continuous interactions: interactions 1-5 included in-between behaviors (such as between behavior 1 and 2 there might be other behaviors), while interaction 6 introduced randomness of motion. This design enhanced the diversity of robot behaviors and facilitated the exploration of human perceptions. However, the continuous interaction approach made it challenging to differentiate between behaviors or classify robot behaviors at any given moment into specific types to understand how they affected human perceptions. Future studies should investigate how human perceptions of robot navigation vary under different robot behaviors. To overcome this limitation, experiments should engage participants with each behavior type individually, allowing for meaningful insights and comparisons.

Contrary to the findings of Koay et al. [356], which reported discomfort among subjects when a robot approached within 3 meters, and in contrast to the principles of proxemics applied in socially aware navigation [12, 325], participants in this study only reported 5 out of 139 instances of the robot being “Too close". This discrepancy could be attributed to the relatively narrow space in the experimental setting (path width ranging from 0.8m to 1.2m, as depicted in Fig. 5.1), which may cause individuals to tolerate closer proximity due to the necessity of passing by the robot [298]. Furthermore, Fig. 5.5 shows that the participants preferred the robot to “Slow down" over move “Away" even when they felt unsafe, which might further indicate an acceptance of closer proximity in narrow spaces. These findings indicate that proxemics rules should be adapted to account for path width, enabling robots to make more context-aware decisions in varying environmental conditions.

Participants have also reported various concerns, see Table 5.4. This finding aligns well with the literature understanding and optimizing human safety, comfort, efficiency, legibility, and avoidance of interference [11, 165, 204, 357, 358]. However, there have hardly been studies investigating human concerns about “Disrespect" and increased “Cognitive load", which deserve further attention. Unlike other concerns, the “Safety" concern was majorly triggered by inappropriate accelerations (see Fig. 5.5). Since safety is the fundamental concern in socially aware navigation, understanding how robot accelerations influence human perceptions and concerns deserves further investigation.

5.4.2. Human-preferred robot adaptations

Unlike previous studies that assumed cooperative navigation between humans and robots [280, 289], this study reveals human preferences for robots to make adaptations while they continue their current behaviors, at least when the robot behaves inappropriately. This is indicated by the participants predominantly reporting solely on robot adaptations without mentioning their own, although Q2.1 openly asked about the interaction as a whole, see Table 5.2. Some participants explicitly mentioned that they would prefer to continue their current behaviors while the robot makes adaptations. Therefore, it is suggested that when a robot is perceived as behaving inappropriately, it should change its own behavior while assuming that humans continue their previous behaviors.

The participants rated the robot adaptations significantly better than the robot continuing its inappropriate behavior (see Table 5.7), which emphasizes the necessity of detecting the perceived appropriateness of robot navigation behavior, thus paving ways for making adaptations [10]. Among all adaptations, the robot moving “Aside, stop” was identified as the most preferred adaptation and was statistically different from all other adaptations (Bonferroni-corrected p-values below the adjusted significance level of $= 4.5 \times 10^{-3}$, see Section 5.3.2). This might indicate that robots could always move “Aside, stop” when perceived as inappropriate by humans, especially in narrow spaces. This experiment was conducted in a relatively narrow space, which might explain why the participants overwhelmingly preferred the robot to move “Aside, stop", and as such, its applicability might not apply to open spaces. Though also in narrow spaces, the study by Senft et al. [298] discovered a human preference for the robots to rotate their bodies rather than a sliding motion. It would be interesting for future studies to also take the robot body rotation into account when comparing different adaptations. Note that the evaluations of the 12 adaptations relied on participants imagining hypothetical scenarios where they assessed the appropriateness of each adaptation. Although this yielded useful insights for designing robot adaptations, it did not fully capture real-world conditions. Future research could benefit from employing the “think aloud" approach, which would allow participants to express their perceived appropriateness and preferred adaptations in real-time [355]. At the same time, the robot could make specific adaptations based on these real-time inputs, offering a more accurate representation of their practical utility. The findings of this study rely on the assumption that robots could detect the perceived appropriateness of their navigation behaviors. While advancements have been made [10, 342], further research is necessary to improve the detection performance, especially in identifying and classifying various inappropriate behaviors/errors. As human responses to different robot errors show considerable variation during human-robot interactions [64], it would be beneficial to investigate how robots can differentiate between types of inappropriate behaviors and adjust their responses for improved performance accordingly.

5.4.3. Additional insights

Most participants (11 out of 12) agreed on a generic adaptation approach for resolving most inappropriate robot behaviors. However, their preference underwent a notable shift when presented with a broader selection of 12 specific adaptations later on. While "Stop" emerged as the most favored choice when participants were asked to propose solutions (8 out of 11), the introduction of the 12 adaptations later on revealed considerable variability in their preferences. Despite this variation, all 11 participants agreed on the importance of the robot moving away rather than slowing down. This preference was further underscored by participants' overwhelmingly positive evaluations of the robot "Aside" compared to merely "Stop" or "Slow down" (see Table 5.7), while the combination of the two-"Aside, stop"-received the highest average rating. This finding emphasizes the significance of the robot creating space for humans after behaving inappropriately while slowing down further contributes to the appropriateness of the adaptation.

Of the 12 participants, 11 preferred the inclusion of external interfaces for robot behavior adaptations. Among various external interfaces considered, 8 participants favored speech or lights, while 2 participants expressed a preference for facial expressions. This underscores the importance of the robot utilizing external interfaces alongside its motion to adapt inappropriate navigation behaviors.

5.5. Conclusions

This study investigated human-preferred adaptations for inappropriate robot navigation behaviors by conducting a series of human-robot interactions, a semi-structured interview, and an online survey. Firstly, 12 participants interacted with a mobile robot in the designed lab narrow setting and identified robot behaviors they perceived as inappropriate. Then, a semi-structured interview was conducted to identify humans' perceived inappropriate robot behaviors and the preferred adaptations. Additionally, all participants were invited back to evaluate the appropriateness of 12 selected robot adaptations for addressing inappropriate robot behaviors. The coding of the interview revealed 139 instances of inappropriate robot behaviors, categorized into 9 distinct types. These were influenced by factors of robot-human relative poses, as well as the robot's speed, acceleration, and motion smoothness. The study also revealed overwhelming negative evaluations of the robot continuing its inappropriate navigation, highlighting the need for the robot to adapt its inappropriate behaviors. As for adaptations, the robot moving to the side of the pathway and stopping was the most preferred and evaluated positively, reflecting its generalizability in addressing inappropriate robot navigation, especially in narrow spaces. Despite the controlled laboratory setting, the small and homogeneous participant sample, and the potential bias by having participants identify inappropriate robot behaviors after all interactions, these conditions were instrumental in allowing the exploration of human perceptions and preferred adaptations. Future work will address these limitations by involving a more diverse participant pool and conducting real-life comparisons of different adaptations to compare their effectiveness in addressing inappropriate robot behaviors.

5.6. APPENDICES

5.6.1. PREFERRED ADAPTATIONS

Table 5.8 presents 34 types of preferred robot adaptations reported by the 12 participants.

Table 5.8: 34 types of preferred robot adaptations

Adaptation	Number	Adaptation	Number
Away	22	Stop	31
Slowly away	6	Quickly stop	11
Quickly away	2	Slowly stop	5
Smoothly away	3	Smoothly stop	3
Perpendicular away	4	Slow down	19
Side	6	Slowly slow down	3
Slowly side	3	Smoothly slow down	2
Quickly side	1	Quickly slow down	1
Retreat	6	Speed up	5
Slowly retreat	3	Slow down, inform	6
Predictable motion	6	Stop, inform	4
Go straight	2	Stop, next to and copy, say sorry	3
Change direction	3	Say sorry, slowly stop	2
Turn right	4	Say sorry, slow down	1
Retreat, stop	9	Slowly slow down, ask for expectations	1
Slow down, change direction	1	Away, inform	2
Side, stop	6	Speech inform action	5

5.6.2. CODING SCHEME

Below is the coding scheme for each question, along with an example provided in Table 5.9.

Table 5.9: Example of coding scheme

Main theme	**Sub-theme**	**Definition**	**Example quote**
Relative pose	Behind, Same	The robot is behind the human, and moving in the same orientation	Video showing the robot moving behind the human with the same motion orientation
Robot behavior factors (Q1.1)	Slow	The robot moves at a slow speed	“The robot moves slowly"
PA (Q1.2)	Acceleration	The robot behaved inappropriately because it made a sudden acceleration	“The robot suddenly accelerated towards me without stopping."
Concern (Q1.3)	Unsafe	The robot’s behavior triggered human concerns about safety	“It’s going to hit me, which makes me feel unsafe."
Preferred adaptation (Q2.1)	Slow down and say sorry	The participant preferred the robot to slow down and apologize	“If the robot can say sorry to me. Action-wise, like slowing down."
Evaluation of the next action (Q2.2)	Slowed down; -2	The robot slowed down after behaving inappropriately	“It slowed down. It’s already about to hit me, so strongly inappropriate (-2)."
Evaluation of the 12 adaptations (Q2.3)	-2	The robot adaptation of “continue" is rated -2	“continue", strongly inappropriate (-2)

5

1. **Robot-human relative pose:** This category systematically analyzes interactions recorded in videos supplemented by interview data. Each interaction is timestamped to facilitate precise identification. The coding encompasses two primary dimensions:
   a) **Relative Position:** Coded relative to the participant’s walking direction and categorized as front, behind, left, or right.
   b) **Relative Orientation:** Coded relative to the participant’s walking direction and categorized as same, opposite, left, right.
2. **Participants’ reports of robot behavior factors (Q4.1.1):** Participants’ descriptions from interviews are also coded, focusing on aspects like robot **Speed** (e.g., “Fast", “Slow", “Stationary"), **Acceleration** (e.g., “Accelerating", “Decelerating"), and motion **Smoothness** (e.g., “Jerky", “Snake-like", “Unexpected").
3. **Inappropriate Robot Navigation Behavior (Q4.1.2):** This category involves coding participants’ perceptions of inappropriate robot behavior. The coding process involves:
   a) Identifying verbs participants use to describe robot behaviors, such as “Accelerated", “Blocked", etc.
   b) Merging semantically similar verbs into unified themes, for example, “accelerated", “Sped up", and “Making accelerations", are grouped under the theme of “Accelerate".
4. **Participants’ Concerns (Q4.1.3):** This category involves coding the participants’ concerns triggered by inappropriate robot behaviors. The coding process

includes:

a) Identifying key terms related to concerns, such as “Collisions", “Delays", “Increase my cognitive load", etc.

b) Grouping related terms into thematic categories, such as “Unsafe", “Collisions", etc., under the broader theme of “Safety".

5. **Participants’ Preferred Robot Adaptations (Q4.2.1):** This category codes the participants’ preferences for robot adaptations after inappropriate robot behaviors. The coding includes:

   a) Coding key terms, particularly verbs and adverbs, representing preferred interactions, such as “Slow down", “Stop", “Aside", etc.

   b) Grouping similar terms into themes, like “Stop", “Zero speed", “Pause", under the theme of “Stop."

   c) Creating sub-categories based on the intensity of the preferred behavior, e.g., “Slowly stop", “Quickly stop", etc.

6. **Participants’ Evaluations of the Robot’s Next Action (Q4.2.2):** This category involves coding the participants’ interpretations of robot actions after the button press and their evaluations of these actions. The process includes:

   a) Coding key terms, mainly verbs, to categorize robot behaviors, such as “Slowed down", “Stopped", etc.

   b) Grouping similar terms under themes like “Stopped", “Paused", etc., under the broader theme of “Stop".

   c) Quantifying the participants’ evaluations on a 5-point scale, ranging from “Strongly inappropriate" (-2) to “Strongly appropriate" (2).

7. **Participant Evaluations of 12 Robot Adaptations (Q4.2.3):** This category involves coding the participants’ ratings of different adaptations. The process includes:

   a) Quantifying the participants’ evaluations on a 5-point scale, ranging from “strongly inappropriate" (-2) to “strongly appropriate" (2).

   b) Aligning the evaluations with its related adaptation and inappropriate robot behavior instance.

6

Conclusion & Discussion

This thesis contributes to socially aware navigation (SAN) by enabling robots to utilize human-perceived appropriateness (PA) information to enhance their interaction quality with humans. The core aim was to enable robots to adapt their behavior based on human feedback.

6.1. Main findings and conclusions

This dissertation investigates how robots can understand and adapt to humans in social environments to improve the appropriateness of their navigation behavior. We address this with the following research questions:

RQ1: In existing research on socially aware navigation, what types of information communicated between humans and robots have been studied, and what social cues have been used for this communication? To understand the exchange of information between humans and robots, we conducted a scoping review in Chapter 2, identifying 176 relevant studies. Results revealed that while a majority of 87% studies focused on how robots process human information, only 13% examine how humans perceive and process robot information. This highlights the need for more research focusing on how to design effective communication strategies for robots to convey information to humans. This could involve utilizing various modalities, such as visual displays, auditory signals, or even haptic feedback, to ensure clear communication.

The exchanged information within socially aware navigation studies were further categorized into three major types: future behaviors (long-term goals, short-term intentions, legibility, etc); social signals (emotion, dominance, activity, group dynamics, PA, etc); and proxemic spaces (intimate, personal, social, public). The quantitative analysis on this information revealed that social signals (compared to future behaviors and proxemics spaces) and especially human PA of robot behavior were understudied. PA is a crucial feedback to enable identification of inappropriate robot behaviors for adaptations. Without understanding PA, robots risk making repeated mistakes, which could significantly hinder their acceptance in social environments. This gap, with so few addressing it (e.g., Vroon et al. [10]), motivates our further investigations on PA.

This study further revealed the detailed relations between social cues and different types of information. Firstly, a thematic analysis was conducted on the collected studies to collect a comprehensive list of social cues specifically relevant to navigation, which expanded upon the cues typically summarized in non-navigational contexts [359]. These include social cues of: head, velocity, position, motion, distance, orientation, posture, gaze, trajectory, facial expressions, speech, gestures, F-formations, gender, and external human-machine interfaces (eHMIs). Secondly, quantitative analysis of the relations identified the distinct roles of various social cues in communicating different information. Orientation, trajectory, and position were found to be crucial for predicting future behavior, enabling robots to both anticipate human behaviors and signal their own intentions. Social signals such as emotions were primarily conveyed through facial expressions, and dominance mostly through trajectories. Proxemic spaces relied majorly on social cues of position and orientation. These findings highlight how a careful selection and combination of social cues can enable both effective robot communication and accurate understanding of human behavior in social navigation.

RQ2: What are the key factors contributing to human yielding to conflicts in public spaces, and how do they influence human PA of robot navigation behaviors? Failures often arise when humans and robots do not have a consensus when it comes to yielding. While significant research has focused on conflict avoidance [338, 350], factors influencing human yielding behaviors are often neglected. To investigate human yielding, we conducted a field observation in Chapter 3, where a robot was manually controlled to ignore humans, thereby triggering conflicts. Based on 427 observed instances of human yielding behaviors, we identified various contributing factors, covering human activities and demographics, robot behavior, and environmental settings. In most cases, humans yielded to the robot, avoiding potential collisions well before they would occur, despite in narrow spaces where they exhibited richer perceptions and responses. This observation aligns with previous research showing that conflicts are more frequent in narrow spaces [298]. In these narrow spaces, humans tended to have a more negative perception of the robot and expected it to adapt to them. This negative perception was further intensified by inappropriate robot behaviors, such as blocking the path or moving too close, which Koay et al. also found to trigger human discomfort [323]. Moreover, such negative perceptions often elicited certain human reactions (social cues, attention shifts, emotional responses). These findings highlight the need for further investigation into how robot errors, reflected through human PA, can be inferred from social cues [10], emotion, and attention.

Furthermore, Chapter 4 identified the influence of robot navigation behaviors on PA based on human-robot interaction experiments. Specific behaviors such as blocking and sudden direction changes correlated with negative PA, while behaviors of remaining stationary and predictable movement at a distance elicited more positive PA. Chapter 5 further elaborated on these findings, identifying specific influencing factors, including robot-human relative poses and robot motion characteristics (speed, acceleration, smoothness), on the perception of robot behavior.

RQ 3: How can we detect the perceived appropriateness of robot navigation behavior in narrow environments? Extending beyond the work of Jered et al. on detecting the PA of robot positioning behavior [10], we conducted eight human-robot interactions (inspired by the work of Koay et al. [323]) to create the PARSNiP dataset for PA detection in Chapter 4. Apart from social cues, the dataset also contains human emotion and attention. Several common machine learning models were applied to the dataset to assess the contributions of various features in PA detection. The results indicate that incorporating emotional and attentional features greatly improves detection performance, consistent with prior work emphasizing the significance of human emotions in identifying perceived robot mistakes in human-robot team-working [318]. Specifically, accuracy improved from 63% to 68% when using algorithm-predicted emotional and attentional data, and further increased to 79% when relying on participant-reported emotion and attention data. Therefore, this study provides a dataset that enables PA detection and offers insights into how emotion, attention, and possibly other social signals, such as dominance levels, could further contribute to PA detection.

RQ4: How do humans prefer a robot to adapt its inappropriate navigation behavior in narrow environments? Human-robot interaction experiments, conducted in a narrow setting and detailed in Chapter 5, triggered a range of inappropriate robot behaviors. Participant interviews then identified preferred adaptations for various PA scenarios. Participants overwhelmingly perceived the robot's continued inappropriate behavior as unacceptable, contrasting with prior research suggesting increased likeability of errant robots [64]. This discrepancy likely stems from contextual differences: unlike non-navigation settings with lower safety risks, navigation errors directly impact participant safety, reducing tolerance for inappropriate behaviors. Furthermore, participants predominantly preferred the robot to move "Aside, stop", regardless of the specific inappropriate behavior, suggesting this as a potential generic adaptation strategy in narrow spaces.

6.2. REFLECTIONS

6.2.1. PA FRAMEWORK

Robots navigating in human environments often exhibit behaviors perceived as inappropriate by humans (e.g., blocking a narrow path, approaching too closely), leading to discomfort, disruption, and potential safety issues. This thesis argues that incorporating human-perceived appropriateness (PA) into robot navigation systems is crucial for creating socially acceptable robots. To structure this investigation, a comprehensive framework for understanding and addressing PA in robot navigation was developed, as illustrated in Figure 6.1.

This PA framework provides a holistic view of the factors, responses, detection methods, and adaptation strategies relevant to achieving socially acceptable robot

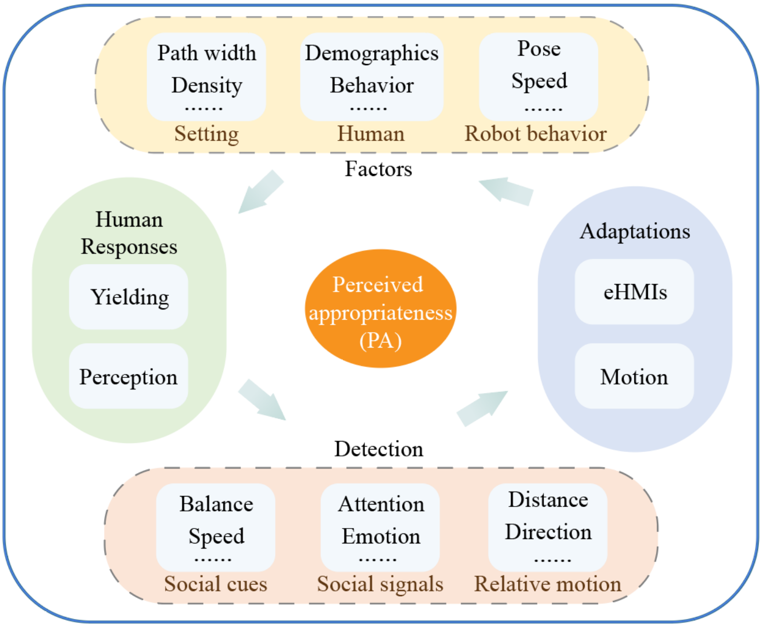


Figure 6.1: The Perceived Appropriateness (PA) Framework

navigation. At its core is the concept of Perceived Appropriateness (PA). This perception is influenced by a set of **Factors**, categorized into three main groups: *Setting*, which encompasses environmental characteristics like path width and environmental density, etc; *Human* factors, including demographics such as age and gender, and human behavior, etc; and *Robot behavior*, encompassing the robot's pose and navigation speed, etc. These influencing factors directly impact **Human Responses**, which manifest in two primary ways: *Yielding behaviors*, such as changing trajectory or speed to avoid the robot, and *Perception*, the subjective evaluation of the robot's behavior. Understanding these responses is crucial for PA **Detection**. The framework highlights three key categories of features: *Social cues*, like a person's balance and walking speed; *Social signals*, including a person's attention level and emotion; and *Relative motion* features, such as the distance and relative direction between the human and the robot. Finally, the framework addresses **Adaptations**, the strategies a robot can employ to improve its perceived appropriateness. These adaptations are divided into two main categories: adjustments of the robot's *Motion*, such as changing its speed or trajectory, and the use of external *Human-Machine Interfaces (eHMIs)*, such as lights or displays, to communicate with humans. The framework posits a cyclical relationship: factors influence human responses, which are detected, leading to adaptations that, in turn, affect the influencing factors, continuously shaping the PA of the robot's behavior.

It's important to acknowledge that the specific factors influencing PA, and the effectiveness of different adaptation strategies, can vary significantly across different contexts. This framework, therefore, is presented as a flexible structure that can be adapted and tailored. Several key dimensions of variability warrant consideration:

Cultural Contexts: The perception of appropriate robot behavior is heavily influenced by cultural norms. For example, personal space preferences vary considerably between cultures [38]. What is considered an acceptable approach distance in one culture might be perceived as intrusive or aggressive in another. The framework's *Human Factors* (specifically, aspects of demographics) must be tailored to the specific cultural context in which the robot works. This might involve training models on culturally-specific datasets or incorporating explicit rules derived from sociological research.

Robot Type and Application: The function and form of the robot itself significantly impact how the framework is applied. The relevant factors influencing PA might differ. For a service robot, pedestrian density and obstacles are paramount. For an industrial robot, task-specific factors (tool use, operational efficiency) and the human worker's role become critical. For a healthcare robot, human factors like patient vulnerability and emotional state are key. Consequently, human responses also differ, ranging from yielding behaviors in public spaces to direct instructions in industrial settings, or subtle cues of discomfort in healthcare. Detection methods must be tailored accordingly. Service robots might rely on visual cues and proxemics, industrial robots on precise sensor data of tools, and healthcare robots on physiological signals. Finally, appropriate adaptation strategies might also vary according to different robot types and applications.

Setting: The physical and social characteristics of the setting itself exert a profound influence on PA. A narrow hallway, for example, inherently limits maneuverability and increases the likelihood of close encounters, amplifying the importance of factors like approach speed and trajectory. In contrast, a wide-open plaza allows for greater separation distances, potentially reducing the perceived severity of minor navigational errors. The setting's social context also matters; a robot navigating a busy train station during rush hour will face different expectations and tolerances than one operating in a quiet library. Crowd density, the presence of static obstacles (e.g., benches, kiosks), and even ambient noise levels can all shape human perceptions. Thus, a thorough understanding of the setting's physical and social attributes is crucial for effectively applying the PA framework.

6.2.2. THESIS REFLECTION

Based on the framework, this thesis investigates the four key aspects of PA, especially in narrow environments rich in conflicts and negative human perceptions: 1) the factors influencing human yielding behavior and their perceptions of robot behavior appropriateness, 2) human responses to PA, 3) the detection of PA, and 4) human-preferred robot adaptations.

Factors This thesis provided empirical insights into human yielding behaviors in response to conflicts with a robot and factors contributing to such yielding in Chapter 3. The yielding behaviors examined included yielding timing, changes in navigation (speed/trajectory), and bodily responses. Results revealed a human preference to yield early, highlighting the importance of robots minimizing interference to humans, in line with findings from Peddi et al [347]. This might also explain why humans preferred to change their trajectories instead of speed in most yielding instances, a strategy which could enable humans to avoid conflicts as soon as possible. These yielding behaviors were influenced by factors including robot-human relative motion (crossing, head-on, or from behind), the environmental setting (path width, environmental density), and human states such as whether the pedestrian was stationary or moving, the pedestrian's assumed age and gender, their engagement level (distracted or not) and activities (cycling, carrying items, or using strollers). For instance, a stationary pedestrian talking on the phone was far less likely to yield (5 out of 14 such encounters led to an emergency stop) compared to a young pedestrian strolling around. However, yielding itself did not necessarily indicate a negative perception of the robot; interviews revealed that while many pedestrians yielded, they often described the encounter as "interesting" or "strange," suggesting curiosity rather than disapproval. It was in narrow environments, where maneuverability was limited [298], that pedestrians more strongly expressed the need for the robot to adapt its behavior, often giving the robot less space and gazing at it longer and sometimes with emotional reactions. These behaviors hint at the possibility of using such cues and responses for PA detection. Beyond path width, Chapter 4 also revealed the influence of robot navigation behaviors on human perceptions of appropriateness (PA). Behaviors such as blocking the path and sudden changes in direction correlated with negative PA, whereas remaining stationary and making predictable movements at a distance elicited positive PA.

Responses Individual responses to the robot varied significantly, as evidenced by differing reactions even to the same robot behaviors (see Chapter 4). Some participants exhibited discomfort or disorientation, while others maintained a positive attitude regardless of the robot's behaviors. This range of responses, from completely inappropriate to completely appropriate, highlights the complexity of human-robot interactions and the subjective nature of PA. The specifics of how robot behavior parameters affected PA were further detailed in Chapter 5, including the influence of relative robot-human poses and robot motion characteristics (e.g., speed and acceleration). Specifically, a robot positioned directly in front of or behind the human triggered the most inappropriate behavior compared to a side positioning. A robot moving in parallel (in the same direction) as the human, but accelerating, was also frequently perceived as inappropriate. Furthermore, "jerky" or "unpredictable" robot movements (related to motion smoothness) often led to perceptions of inappropriateness. These results indicate a complex interplay between robot motion and human perception, and that researchers and designers can strategically manipulate these parameters to influence human perceptions and yielding behavior, ultimately shaping desired interactions.

Detection This thesis demonstrates that emotion and attention play a more significant role in improving PA detection performance than social cues or motion features. These findings corroborate prior research, such as that by Loureiro et al., which emphasizes the role of emotion in identifying perceived robot errors [318]. This significant contribution of emotion and attention likely stems from their roles as indicators of human comfort and perceived safety; negative emotions and heightened or fluctuating attention may signal that the robot's behavior is disruptive, unpredictable, or even threatening, thus leading to lower PA ratings. Serving also as important indicators, other social signals such as dominance levels [52] and attitudes [360] might also contribute to PA detection and deserve further investigation. Beyond overall appropriateness, our findings highlight that humans also form specific perceptions about the robot's behavior, such as "blocking", "squeezing", "sudden motion", "threatening", "unnatural", "fast", etc. These nuanced perceptions are crucial and should be further investigated. This understanding could inform researchers to develop datasets that enable robots not just to detect whether behavior is appropriate, but also to identify the specific type of inappropriate perception (e.g., blocking vs. squeezing), leading to more targeted and effective behavioral adaptations.

Adaptation Chapter 5 revealed a predominantly negative human perception of the robot's continued inappropriate behavior, underscoring the necessity for robots to adapt such behaviors. The identified adaptations included simple actions like stopping ("Stop"), slowing down ("Slow"), and moving away from the participant ("Away"). More complex adaptations involved combinations of movement and communication, such as slowing down and verbally informing the participant of the intended action ("Slow, speak"), or stopping and using a visual cue (e.g., a light) to indicate a change in plan ("Stop, light"). A comparison of these adaptations revealed a strong preference for "Aside, stop" across a variety of inappropriate behaviors, suggesting its potential as a general-purpose adaptation strategy, particularly in narrow spaces. This preference likely stems from the combination of immediate risk mitigation (stopping) and clear communication of intent (moving aside), providing the participant with both physical safety and a predictable, unobstructed path. While the PA framework (Figure 6.1 includes external Human-Machine Interfaces (eHMIs) as a potential adaptation strategy, this thesis primarily focused on motion-based adaptations. Further research is needed to fully explore the role and effectiveness of various eHMIs in improving perceived appropriateness in different navigational contexts.

6.3. Implications and implementations

This section explores the practical implications of the findings presented in this thesis, offering guidance for designing robot navigation behaviors that are not only functional but also socially intelligent and considerate of human perceptions and expectations.

Chapter 2 revealed the understudied nature of social signals, particularly real-time feedback signals such as human PA of robot behavior. This highlights a critical need for robots to move beyond simply processing what information humans are communicating (e.g., trajectory, intended goal) to also understanding how that information is being received and perceived. When designing robots, researchers should carefully consider which type of human information is most relevant to the robot's tasks and how this information can be inferred from specific social cues. For example, a robot designed primarily for navigation in crowded spaces might prioritize understanding future states (predicting pedestrian trajectories) and social signals (detecting discomfort or frustration). The review also provided a detailed mapping between information types and the selection of appropriate social cues. This mapping can guide the selection of appropriate sensors. For instance, predicting future human behavior often relies on cues like head orientation, trajectory, position, and motion direction [361], suggesting the use of cameras and potentially depth sensors [362]. Detecting social signals like emotion, however, might require facial expression analysis [281], necessitating high-resolution cameras, and potentially even physiological sensors in the future. Critically, designers should recognize the interconnectedness of information types. Specifically, when designing a robot that can detect human social signals, it's crucial to consider how these signals inform the processing of other information, such as prediction and proxemics zones. This makes the robot more informative. A robot should also not process future states, social signals, and proxemic spaces in isolation [85, 363]. Instead, it should integrate these different information streams to form a holistic understanding of the situation. If, for example, a robot is predicting a human approaching—despite being within the human's personal space—and detects the human showing positive emotions, the robot can infer that the human is interested in interaction and adjust its behavior accordingly.

Chapter 3 revealed a human preference to avoid conflicts with robots from early on, implying that robot behavior should be designed for proactive and legible conflict avoidance [350], rather than relying solely on reactive responses. Robots should not wait until a conflict is imminent before taking action. This requires not only accurate prediction of human motion but also clear communication of the robot's own intentions. Legibility, as discussed in Chapter 2, can be achieved through careful motion planning (e.g., predictable trajectories, smooth transitions) and the use of eHMIs (e.g., directional indicators, lights, displays)[161]. Additionally, the study identified several factors influencing yielding behaviors, including environmental settings, human demographics and activities, and robot behavior. These findings suggest that robot navigation should be context-aware. A robot should adapt its behavior based on the specific environment (e.g., yielding more readily in a narrow hallway than in a wide-open space), the characteristics of the humans it encounters (e.g., exhibiting greater deference to an elderly person or someone carrying a large object), and its own actions (avoiding sudden movements or close approaches).

Chapter 4 demonstrated that incorporating human emotion and attention significantly improves the accuracy of PA detection. This finding underscores the need for robot behavior design to extend beyond purely kinematic data (position, velocity, trajectory). Robots should be equipped with sensors and algorithms capable of inferring and responding to human social signals, particularly emotion and attention. This includes, specifically, facial expressions, gaze, vocal intonation, and body language [43]. The significant improvement in PA detection when using reported emotion and attention (compared to predicted data) highlights the potential benefits of accurate social signal detection. This motivates continued research into advanced emotion recognition and attention estimation algorithms, which would not only enhance the robot's understanding of human emotional states, but by extension improve PA detection accuracy, contributing to fewer social errors [364].

Chapter 5 revealed a strong human preference for the "Aside, stop" adaptation (moving to the side of the pathway and stopping), suggesting its potential as a general-purpose strategy, particularly in narrow environments. However, it's also important to recognize that the "best" adaptation will likely depend on the specific context. A robot should be capable of selecting from a repertoire of adaptations, considering the type of inappropriate behavior, the environment, the human's emotional state, and its own capabilities. Furthermore, robot adaptation should not be viewed as a one-time action, but rather an ongoing process of monitoring and adjustment. The robot should continuously monitor human responses (through social cues and other signals) and adjust its behavior accordingly. This iterative process of perception, adaptation, and re-evaluation is crucial for achieving truly socially acceptable navigation.

6.4. LIMITATIONS AND FUTURE RESEARCH

This thesis has investigated perceived appropriateness (PA) primarily within the context of robot navigation in narrow environments, exploring the interplay of various factors, human responses, detection methods, and adaptation strategies. While these findings provide valuable insights, it is crucial to acknowledge the limitations of this work and discuss its broader implications for robot design and human-robot interaction.

6.4.1. SCOPE AND CONTEXTUAL FACTORS

The focus on narrow environments, while intentional to highlight the importance of PA in situations with limited maneuverability and frequent potential conflicts, inherently limits the direct generalizability of the findings to wider, less-restricted spaces. Open spaces, while still relevant to PA, present different dynamics. A robot crossing a large plaza, for instance, might disrupt the flow of many pedestrians, leading to negative PA even if collisions are easily avoided. Similarly, a robot positioned near a conversing group might be perceived as intrusive, even if maintaining a technically "safe" distance. These examples illustrate that PA in open spaces is often tied to more subtle violations of social norms, not solely collision avoidance, which deserves further investigations.

Furthermore, the specific robot behaviors investigated, while informed by existing literature and pilot studies, represent only a subset of potential inappropriate behaviors. Real-world scenarios are far more complex and nuanced, involving a wider range of human activities (conversations, queuing, carrying objects), group dynamics (friends, families walking together), and environmental factors (crowd density, ambient noise). Future research must broaden the scope to encompass these complexities, exploring how PA is shaped by diverse social contexts and individual differences.

The intentional use of non-humanoid robots (Clearpath Husky and Jackal) in this research allowed for a focused investigation of navigation behavior without the confounding effects of anthropomorphic features. However, this design choice limits the direct applicability of the findings to humanoid robots. People tend to attribute greater social awareness and intentionality to robots that resemble humans, leading to different expectations and interpretations of their behavior. A slight trajectory deviation, easily dismissed as a minor navigational error in a non-humanoid robot, might be perceived as rude or aggressive in a humanoid robot. Therefore, research involving humanoid robots must carefully consider the impact of appearance, facial expressions, gaze, posture, and even tone of voice on PA.

6.4.2. Human responses and measurement

This thesis primarily examined human responses to PA through yielding behaviors, gaze patterns, and self-reported perceptions. While these measures provide valuable data, they represent only a partial picture of the complex human response to inappropriate robot behavior. Future research should incorporate a wider range of measures, including verbal reactions (complaints, comments), facial expressions, physiological data (heart rate variability, skin conductance, brain activity), and even subtle shifts in body language. There could be a time delay between robot behavior and human reactions. Moreover, these responses may be dynamic, changing from momentary confusion to increased anxiety or discomfort. Capturing these nuances requires sophisticated methods that combine self-report techniques (questionnaires, interviews, "think-aloud" protocols) with objective physiological and behavioral measures.

6.4.3. PA detection and adaptation

While this thesis demonstrated the promise of using machine learning to detect PA, particularly by incorporating emotional and attentional features, significant challenges remain for real-world deployment. The reliance on retrospective self-reports of emotion and attention in the dataset creation introduces potential biases. Future research should explore real-time emotion recognition algorithms and integrate a broader range of social cues, including facial expressions, body posture, and vocal intonation. Furthermore, temporal models, such as recurrent neural networks or Hidden Markov Models, are needed to capture the dynamic nature of PA and how it evolves over time.

The finding that "Aside, stop" was a generally preferred adaptation in narrow spaces provides a valuable starting point, but it is unlikely to be a universally optimal solution. Different types of inappropriate behavior, varying environmental constraints, and individual preferences will necessitate a wider repertoire of adaptation strategies. Future research should investigate the effectiveness of multi-modal adaptations, combining kinematic adjustments (changes in trajectory, speed) with communication methods (facial expressions, gaze, posture, or external human-machine interfaces (eHMIs) like lights, displays, or even speech). The selection of appropriate adaptations should be context-aware, taking into account the severity of the error, the user's emotional state, and the robot's own capabilities. For instance, a robot that has committed a minor navigational error might simply adjust its trajectory and offer a brief apology through a visual display. A more serious error, such as nearly colliding with a person, might require a more substantial adaptation, such as stopping completely, apologizing verbally, and waiting for the person to pass.

6.4.4. Broader implications

Beyond the specific limitations discussed above, this research raises broader questions about the long-term implications of integrating robots into social environments. As robots become more commonplace, human expectations and perceptions of appropriate behavior are likely to evolve. Longitudinal studies are needed to track these changes and to understand how repeated interactions with adaptive robots shape trust, comfort, and overall acceptance. Initially, interactions with robots might be dominated by novelty and a high tolerance for errors. People might be forgiving of awkward movements or misinterpretations of social cues, viewing them as curiosities. However, as familiarity increases, expectations will likely rise. Behaviors once considered minor inconveniences might become sources of significant frustration. For instance, a robot that occasionally blocks a hallway might be initially amusing, but if it persistently obstructs the path, it will quickly be perceived as highly inappropriate. Conversely, positive, consistent adaptation to PA could foster trust and acceptance. If a robot consistently demonstrates an understanding of personal space, anticipates human needs, and gracefully corrects its errors, humans may become more comfortable interacting with and relying on robots in various contexts.

Furthermore, while this thesis focused primarily on robot navigation, the core concept of PA extends to other areas of human-robot interaction. Whether a robot is assisting with a task, providing information, or simply coexisting in a shared space, its actions should align with human expectations and social norms. The specific factors influencing PA, the relevant human responses, and the appropriate adaptation strategies will vary depending on the context and the task, but the fundamental principle of designing for perceived appropriateness remains crucial. Consider these non-navigation scenarios: A collaborative assembly robot working alongside a human on an assembly line must not only perform its tasks efficiently but also in a way that is perceived as appropriate; if the robot consistently reaches across the human's workspace, creating a feeling of encroachment or risk, this would be a violation of PA, even if no physical contact occurs, requiring the robot to adapt its movements and timing to respect the human's personal space and workflow. Similarly, a receptionist robot providing information must interact in a socially appropriate manner; consistently misinterpreting user requests, providing irrelevant information, or using an inappropriate tone of voice would lead to it being perceived as unhelpful and frustrating, even if it technically provides the correct answer eventually, necessitating adaptations such as rephrasing responses, offering alternative solutions, or escalating the interaction to a human operator. An elderly-care robot offering companionship also needs to understand social appropriateness, such as giving the elderly personal space if required; otherwise the elderly might feel uncomfortable.

Future research should also consider the ethical implications of PA-aware robots. As robots become more adept at understanding and responding to human emotions and social cues, there is a potential for manipulation or exploitation. It is essential to develop ethical guidelines and design principles that ensure robots are used to enhance human well-being and promote positive social interactions, rather than to deceive or control.

In conclusion, this thesis represents a significant step towards understanding and addressing inappropriate robot behaviors. By systematically investigating the factors influencing PA, exploring human responses, developing detection methods, and evaluating adaptation strategies, this work lays the foundation for creating robots that are not only functional and efficient but also socially intelligent and accepted members of our human environments. Continued research in this area is essential to realize the full potential of social robotics and to ensure a harmonious coexistence between humans and robots in the future.

Bibliography

[1] E. Huang. "In China, a robot has started delivering packages to people". In: *Retrieved July* 2 (2017).

[2] A. Meghdari, A. Shariati, M. Alemi, A. A. Nobaveh, M. Khamooshi, and B. Mozaffari. "Design performance characteristics of a social robot companion "Arash" for pediatric hospitals". In: *International Journal of Humanoid Robotics* 15.05 (2018), p. 1850019.

[3] G. Nejat and M. Ficocelli. "Can I be of assistance? The intelligence behind an assistive robot". In: *2008 IEEE International Conference on Robotics and Automation*. IEEE. 2008, pp. 3564–3569.

[4] C. Nass, J. Steuer, and E. R. Tauber. "Computers are social actors". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1994, pp. 72–78.

[5] B. J. Fogg. "Persuasive technology: using computers to change what we think and do". In: *Ubiquity* 2002.December (2002), p. 2.

[6] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch. "Human-aware robot navigation: A survey". In: *Robotics and Autonomous Systems* 61.12 (2013), pp. 1726–1743.

[7] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard. "Socially compliant mobile robot navigation via inverse reinforcement learning". In: *The International Journal of Robotics Research* 35.11 (2016), pp. 1289–1307.

[8] M. Shiomi, F. Zanlungo, K. Hayashi, and T. Kanda. "Towards a socially acceptable collision avoidance for a mobile robot navigating among pedestrians using a pedestrian model". In: *International Journal of Social Robotics* 6.3 (2014), pp. 443–455.

[9] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh. "Core challenges of social robot navigation: A survey". In: *arXiv preprint arXiv:2103.05668* (2021).

[10] J. Vroon, G. Englebienne, and V. Evers. "Detecting Perceived Appropriateness of a Robot's Social Positioning Behavior from Non-Verbal Cues". In: *2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI)*. IEEE. 2019, pp. 216–225.

[11] Y. Gao and C.-M. Huang. "Evaluation of socially-aware robot navigation". In: *Frontiers in Robotics and AI* 8 (2022), p. 721317.

[12] J. Rios-Martinez, A. Spalanzani, and C. Laugier. “From proxemics theory to socially-aware navigation: A survey”. In: *International Journal of Social Robotics* 7 (2015), pp. 137–153.

[13] J. Lunenburg, R. van de Molengraft, and M. Steinbuch. “A representation method based on the probability of collision for safe robot navigation in domestic environments”. In: *Autonomous Robots* 42 (2018), pp. 601–614.

[14] J. Gines Clavero, F. Martín Rico, F. J. Rodríguez-Lera, J. M. Guerrero Hernández, and V. Matellán Olivera. “Impact of decision-making system in social navigation”. In: *Multimedia Tools and Applications* 81.3 (2022), pp. 3459–3481.

[15] J. Oh, J. Heo, J. Lee, G. Lee, M. Kang, J. Park, and S. Oh. “SCAN: Socially-Aware Navigation Using Monte Carlo Tree Search”. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2023, pp. 7576–7582.

[16] J. Lian, W. Yu, K. Xiao, F. Qu, and C. Liu. “Social Robot Navigation Based on a 2D Gauss-Gumbel Spatial Density Model in Human-Populated Environments”. In: *The International Conference on Image, Vision and Intelligent Systems (ICIVIS 2021)*. Springer. 2022, pp. 1133–1143.

[17] N. J. Hetherington, E. A. Croft, and H. M. Van der Loos. “Hey robot, which way are you going? nonverbal motion legibility cues for human-robot spatial interaction”. In: *IEEE Robotics and Automation Letters* 6.3 (2021), pp. 5010–5015.

[18] S. Gopalakrishnan and S. Kambhampati. “Minimizing robot navigation-graph for position-based predictability by humans”. In: *arXiv preprint arXiv:2010.15255* (2020).

[19] C. Bothe, F. Garcia, A. Cruz Maya, A. K. Pandey, and S. Wermter. “Towards dialogue-based navigation with multivariate adaptation driven by intention and politeness for social robots”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11357 LNAI (2018), pp. 230–240. ISSN: 16113349. DOI: 10.1007/978-3-030-05204-1_23.

[20] S. M. Fiore, T. J. Wiltshire, E. J. Lobato, F. G. Jentsch, W. H. Huang, and B. Axelrod. “Toward understanding social cues and signals in human–robot interaction: effects of robot gaze and proxemic behavior”. In: *Frontiers in psychology* 4 (2013), p. 859.

[21] A. Alahi, V. Ramanathan, and L. Fei-Fei. “Socially-aware large-scale crowd forecasting”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 2203–2210.

[22] M. Moussaid, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz. “The walking behaviour of pedestrian social groups and its impact on crowd dynamics”. In: *PloS one* 5.4 (2010), e10047.

[23] H. Xue, D. Q. Huynh, and M. Reynolds. "SS-LSTM: A hierarchical LSTM model for pedestrian trajectory prediction". In: *2018 IEEE winter conference on applications of computer vision (WACV)*. IEEE. 2018, pp. 1186–1194.

[24] C. Chen, Y. Liu, S. Kreiss, and A. Alahi. "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning". In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 2019, pp. 6015–6022.

[25] M. Kuderer. "Feature-Based Prediction of Trajectories for Socially Compliant Navigation". In: *Robotics* (2019). DOI: 10.7551/mitpress/9816.003.0030.

[26] C. Rösmann, M. Oeljeklaus, F. Hoffmann, and T. Bertram. "Online trajectory prediction and planning for social robot navigation". In: *2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE. 2017, pp. 1255–1260.

[27] S. Kim, S. J. Guy, W. Liu, D. Wilkie, R. W. Lau, M. C. Lin, and D. Manocha. "BRVO: Predicting pedestrian trajectories using velocity-space reasoning". In: *The International Journal of Robotics Research* 34.2 (2015), pp. 201–217.

[28] X.-T. Truong and T. D. Ngo. "Toward socially aware robot navigation in dynamic and crowded environments: A proactive social motion model". In: *IEEE Transactions on Automation Science and Engineering* 14.4 (2017), pp. 1743–1760.

[29] K. Qian, X. Ma, X. Dai, and F. Fang. "Robotic etiquette: Socially acceptable navigation of service robots with human motion pattern learning and prediction". In: *Journal of Bionic Engineering* 7.2 (2010), pp. 150–160.

[30] S. Gopalakrishnan and S. Kambhampati. "Minimizing Robot Navigation Graph for Position-Based Predictability by Humans." In: *AAMAS*. 2022, pp. 1604–1606.

[31] A. V. Taylor, E. Mamantov, and H. Admoni. "Observer-aware legibility for social navigation". In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2022, pp. 1115–1122.

[32] K. C. Jeppesen, L. Bodenhagen, and N. Krüger. "Socially Acceptable Behaviour for Robots Approaching Humans using an Adaptable Personal Space". In: *VISIGRAPP (2: HUCAPP)*. 2017, pp. 116–121.

[33] G. Angelopoulos, A. Rossi, C. Di Napoli, and S. Rossi. "You Are In My Way: Non-verbal Social Cues for Legible Robot Navigation Behaviors". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2022, pp. 657–662.

[34] M. M. Neggers, P. A. Ruijten, R. H. Cuijpers, and W. A. IJsselsteijn. "Effect of Robot Gazing Behavior on Human Comfort and Robot Predictability in Navigation". In: *2022 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO)*. IEEE. 2022, pp. 1–6.

[35] C. S. Jakobowsky, A. M. Abrams, and A. M. Rosenthal-von der Pütten. "Gaze-Cues of Humans and Robots on Pedestrian Ways". In: *International Journal of Social Robotics* (2023), pp. 1–15.

[36] X. Yu, M. Hoggenmueller, and M. Tomitsch. “Your Way Or My Way: Improving Human-Robot Co-Navigation Through Robot Intent and Pedestrian Prediction Visualisations”. In: *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 2023, pp. 211–221.

[37] M. Joosse, M. Lohse, N. V. Berkel, A. Sardar, and V. Evers. “Making appearances: How robots should approach people”. In: *ACM Transactions on Human-Robot Interaction (THRI)* 10.1 (2021), pp. 1–24.

[38] E. T. Hall. *The hidden dimension*. Vol. 609. Anchor, 1966.

[39] P. Patompak, S. Jeong, I. Nilkhamhang, and N. Y. Chong. “Learning proxemics for personalized human–robot social interaction”. In: *International Journal of Social Robotics* 12 (2020), pp. 267–280.

[40] F. Melo and P. Moreno. “Socially Reactive Navigation Models for Mobile Robots”. In: *2022 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. IEEE. 2022, pp. 91–97.

[41] M. Shoji, A. A. Saputra, T. Obo, and N. Kubota. “Perceptual System of a Mobile Robot for Autonomous Adaptation to Environmental Changes”. In: *2023 IEEE International Conference on Fuzzy Systems (FUZZ)*. IEEE. 2023, pp. 1–8.

[42] A. Bera, T. Randhavane, and D. Manocha. “The emotionally intelligent robot: Improving socially-aware human prediction in crowded environments”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2019, pp. 0–0.

[43] A. Bera, T. Randhavane, R. Prinja, K. Kapsaskis, A. Wang, K. Gray, and D. Manocha. “How are You Feeling? Multimodal Emotion Learning for Socially-Assistive Robot Navigation”. In: *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020) (FG)* Fg (2020), pp. 894–901. DOI: 10.1109/FG47880.2020.00141.

[44] W. Jin, P. Salaris, and P. Martinet. “Proactive-Cooperative Navigation in Human-Like Environment for Autonomous Robots”. In: *ICINCO 2020-17th International Conference on Informatics in Control, Automation and Robotics*. 2020.

[45] A. Vega, L. J. Manso, R. Cintas, and P. Núñez. “Planning human-robot interaction for social navigation in crowded environments”. In: *Advances in Physical Agents: Proceedings of the 19th International Workshop of Physical Agents (WAF 2018), November 22-23, 2018, Madrid, Spain*. Springer. 2019, pp. 195–208.

[46] F. Haarslev, W. K. Juel, A. Kollakidou, N. Krüger, and L. Bodenhagen. “Context-aware Social Robot Navigation.” In: *ICINCO*. 2021, pp. 426–433.

[47] A. Vinciarelli, M. Pantic, and H. Bourlard. “Social signal processing: Survey of an emerging domain”. In: *Image and vision computing* 27.12 (2009), pp. 1743–1759.

[48] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder. "Bridging the gap between social animal and unsocial machine: A survey of social signal processing". In: *IEEE Transactions on Affective Computing* 3.1 (2011), pp. 69–87.

[49] M. J. Wubben, D. D. Cremer, and v. E. Dijk. "How emotion communication guides reciprocity: Establishing cooperation through disappointment and anger". In: *Journal of Experimental Social Psychology* 45.4 (2009), pp. 987–990. ISSN: 00221031. DOI: 10.1016/j.jesp.2009.04.010.

[50] J. Ginés, F. Martín, D. Vargas, F. J. Rodríguez, and V. Matellán. "Social navigation in a cognitive architecture using dynamic proxemic zones". In: *Sensors* 19.23 (2019), p. 5189.

[51] M. Quiroz, R. Patiño, J. Diaz-Amado, and Y. Cardinale. "Group emotion detection based on social robot perception". In: *Sensors* 22.10 (2022), p. 3749.

[52] T. Randhavane, A. Bera, E. Kubin, A. Wang, K. Gray, and D. Manocha. "Pedestrian dominance modeling for socially-aware robot navigation". In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 2019, pp. 5621–5628.

[53] J. Li, A. Cuadra, B. Mok, B. Reeves, J. Kaye, and W. Ju. "Communicating dominance in a nonanthropomorphic robot using locomotion". In: *ACM Transactions on Human-Robot Interaction (THRI)* 8.1 (2019), pp. 1–14.

[54] Z. Zhao, X. Qi, Y. Zhao, Z. Liao, W. Wang, and X. Yang. "Dynamic and flexible personal space model and parameter generation method for socially-aware robot navigation". In: *2022 12th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE. 2022, pp. 780–784.

[55] B. P. E. Alvarado Vásquez and F. Matía. "A tour-guide robot: Moving towards interaction with humans". In: *Engineering Applications of Artificial Intelligence* 88.March 2019 (2020), p. 103356. ISSN: 09521976. DOI: 10.1016/j.engappai.2019.103356.

[56] M. Nieuwenhuisen and S. Behnke. "Human-Like Interaction Skills for the Mobile Communication Robot Robotinho". In: *International Journal of Social Robotics* 5.4 (2013), pp. 549–561. ISSN: 18754791. DOI: 10.1007/s12369-013-0206-y.

[57] C. M. Ranieri, G. Nardari, A. H. Pinto, D. C. Tozadore, and R. A. Romero. "LARa: A robotic framework for human-robot interaction on indoor environments". In: *Proceedings - 15th Latin American Robotics Symposium, 6th Brazilian Robotics Symposium and 9th Workshop on Robotics in Education, LARS/SBR/WRE 2018* (2018), pp. 383–389. DOI: 10.1109/LARS/SBR/WRE.2018.00074.

[58] M. Moder and J. Pauli. "Proactive Robot Movements in a Crowd by Predicting and Considering the Social Influence". In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2022, pp. 644–651.

[59] A. J. Sathyamoorthy, U. Patel, T. Guan, and D. Manocha. “Frozone: Freezing-free, pedestrian-friendly navigation in human crowds”. In: *IEEE Robotics and Automation Letters* 5.3 (2020), pp. 4352–4359.

[60] B. Xue, M. Gao, C. Wang, Y. Cheng, and F. Zhou. “Crowd-Aware Socially Compliant Robot Navigation via Deep Reinforcement Learning”. In: *International Journal of Social Robotics* 16.1 (2024), pp. 197–209.

[61] A. Bremers, A. Pabst, M. T. Parreira, and W. Ju. “Using Social Cues to Recognize Task Failures for HRI: A Review of Current Research and Future Directions”. In: *arXiv preprint arXiv:2301.11972* (2023).

[62] M. Giuliani, N. Mirnig, G. Stollnberger, S. Stadler, R. Buchner, and M. Tscheligi. “Systematic analysis of video data from different human–robot interaction studies: a categorization of social signals during error situations”. In: *Frontiers in psychology* 6 (2015), p. 931.

[63] D. E. Cahya and M. Giuliani. “Appropriate Robot Reactions to Erroneous Situations in Human-Robot Collaboration”. In: *Social Robotics: 13th International Conference, ICSR 2021, Singapore, Singapore, November 10–13, 2021, Proceedings 13*. Springer. 2021, pp. 166–177.

[64] N. Mirnig, G. Stollnberger, M. Miksch, S. Stadler, M. Giuliani, and M. Tscheligi. “To err is robot: How humans assess and act toward an erroneous social robot”. In: *Frontiers in Robotics and AI* (2017), p. 21.

[65] E. Law, V. Cai, Q. F. Liu, S. Sasy, J. Goh, A. Blidaru, and D. Kulić. “A wizard-of-oz study of curiosity in human-robot interaction”. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2017, pp. 607–614.

[66] M. Ragni, A. Rudenko, B. Kuhnert, and K. O. Arras. “Errare humanum est: Erroneous robots in human-robot interaction”. In: *2016 25th IEEE International symposium on robot and human interactive communication (RO-MAN)*. IEEE. 2016, pp. 501–506.

[67] D. E. Cahya, R. Ramakrishnan, and M. Giuliani. “Static and temporal differences in social signals between error-free and erroneous situations in human-robot collaboration”. In: *International Conference on Social Robotics*. Springer. 2019, pp. 189–199.

[68] P. Trung, M. Giuliani, M. Miksch, G. Stollnberger, S. Stadler, N. Mirnig, and M. Tscheligi. “Head and shoulders: automatic error detection in human-robot interaction”. In: *Proceedings of the 19th ACM international conference on multimodal interaction*. 2017, pp. 181–188.

[69] M. Stiber, R. H. Taylor, and C.-M. Huang. “On using social signals to enable flexible error-aware hri”. In: *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 2023, pp. 222–230.

6

[70] M. Spitale, M. T. Parreira, M. Stiber, M. Axelsson, N. Kara, G. Kankariya, C.-M. Huang, M. Jung, W. Ju, and H. Gunes. "ERR@ HRI 2024 Challenge: Multimodal Detection of Errors and Failures in Human-Robot Interactions". In: *arXiv preprint arXiv:2407.06094* (2024).

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

[72] T. P. Spexard, M. Hanheide, S. Li, B. Wrede, *et al.* "Oops, something is wrong-error detection and recovery for advanced human-robot-interaction". In: (2008).

[73] J. Constine. *Postmates unveils Serve, a friendlier autonomous delivery robot*. 2018. URL: https://techcrunch.com/2018/12/13/postmates-robot-serve/?guccounter=1.

[74] J. Martines. *Pitt suspends delivery robots after wheelchair user reports safety hazard*. 2019. URL: https://triblive.com/local/pittsburgh-allegheny/pitt-suspends-delivery-robots-after-wheelchair-user-reports-safety-hazard/.

[75] R. Kitagawa, Y. Liu, and T. Kanda. "Human-inspired Motion Planning for Omni-Directional Social Robots". In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 2021, pp. 34–42.

[76] C. Jiang, Z. Ni, Y. Guo, and H. He. "Learning human–robot interaction for robot-assisted pedestrian flow optimization". In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 49.4 (2017), pp. 797–813.

[77] G. Ferrer, A. G. Zulueta, F. H. Cotarelo, and A. Sanfeliu. "Robot social-aware navigation framework to accompany people walking side-by-side". In: *Autonomous robots* 41.4 (2017), pp. 775–793.

[78] G. Ferrer, A. Garrell, and A. Sanfeliu. "Robot companion: A social-force based approach with human awareness-navigation in crowded environments". In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2013, pp. 1688–1694.

[79] R. Mirsky, X. Xiao, J. Hart, and P. Stone. "Prevention and Resolution of Conflicts in Social Navigation–a Survey". In: *arXiv preprint arXiv:2106.12113* (2020).

[80] H.-M. Gross, A. Scheidig, K. Debes, E. Einhorn, M. Eisenbach, S. Mueller, T. Schmiedel, T. Q. Trinh, C. Weinrich, T. Wengefeld, *et al.* "ROREAS: robot coach for walking and orientation training in clinical post-stroke rehabilitation—prototype implementation and evaluation in field trials". In: *Autonomous Robots* 41 (2017), pp. 679–698.

[81] M. Poeting, S. Schaudt, and U. Clausen. "A comprehensive case study in last-mile delivery concepts for parcel robots". In: IEEE, 2019, pp. 1779–1788. ISBN: 1728132835.

6

[82] T. Taipalus and K. Kosuge. “Development of service robot for fetching objects in home environment”. In: IEEE, 2005, pp. 451–456. ISBN: 0780393554.

[83] J. Forlizzi and C. DiSalvo. “Service robots in the domestic environment: a study of the roomba vacuum in the home”. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 2006, pp. 258–265.

[84] Y. Terzioğlu, B. Mutlu, and E. Şahin. “Designing social cues for collaborative robots: the role of gaze and breathing in human-robot collaboration”. In: *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*. 2020, pp. 343–357.

[85] G. Barquero, J. Núnez, S. Escalera, Z. Xu, W.-W. Tu, I. Guyon, and C. Palmero. “Didn’t see that coming: a survey on non-verbal social human behavior forecasting”. In: *Understanding Social Behavior in Dyadic and Small Group Interactions*. PMLR. 2022, pp. 139–178.

[86] M. Pascher, U. Gruenefeld, S. Schneegass, and J. Gerken. “How to Communicate Robot Motion Intent: A Scoping Review”. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–17.

[87] R. Möller, A. Furnari, S. Battiato, A. Härmä, and G. M. Farinella. “A survey on human-aware robot navigation”. In: *Robotics and Autonomous Systems* 145 (2021), p. 103837.

[88] V. Braun and V. Clarke. “Using thematic analysis in psychology Qualitative Research in Psychology, 3 (2), 77-101”. In: *Online im Internet* (2006), pp. 1–42.

[89] K. Charalampous, I. Kostavelis, and A. Gasteratos. “Recent trends in social aware robot navigation: A survey”. In: *Robotics and Autonomous Systems* 93 (2017), pp. 85–104.

[90] H. C. Yen, H. P. Huang, and S. Y. Chung. “Goal-directed pedestrian model for long-term motion prediction with application to robot motion planning”. In: *2008 IEEE Workshop on Advanced robotics and Its Social Impacts*. IEEE. 2008, pp. 1–6.

[91] K. Qian, X. Ma, X. Dai, F. Fang, and B. Zhou. “Decision-theoretical navigation of service robots using POMDPs with human-robot co-occurrence prediction”. In: *International Journal of Advanced Robotic Systems* 10.2 (2013), p. 143.

[92] K. Qian, X. Ma, X. Dai, and F. Fang. “Socially acceptable pre-collision safety strategies for human-compliant navigation of service robots”. In: *Advanced Robotics* 24.13 (2010), pp. 1813–1840. ISSN: 01691864. DOI: 10.1163/016918610X527176.

[93] I. Kostavelis, A. Kargakos, D. Giakoumis, and D. Tzovaras. “Robot’s workspace enhancement with dynamic human presence for socially-aware navigation”. In: *international conference on computer vision systems*. Springer. 2017, pp. 279–288.

[94] E. Repiso, A. Garrell, and A. Sanfeliu. “Robot approaching and engaging people in a human-robot companion framework”. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2018, pp. 8200–8205.

6

[95] E. Repiso, A. Garrell, and A. Sanfeliu. “People’s adaptive side-by-side model evolved to accompany groups of people by social robots”. In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 2387–2394.

[96] E. Repiso, A. Garrell, and A. Sanfeliu. “Adaptive Social Planner to Accompany People in Real-Life Dynamic Environments”. In: *International Journal of Social Robotics* (2022), pp. 1–33.

[97] A. F. Foka and P. E. Trahanias. “Probabilistic autonomous robot navigation in dynamic environments with human motion prediction”. In: *International Journal of Social Robotics* 2 (2010), pp. 79–94.

[98] I. Pérez-Hurtado, J. Capitán, F. Caballero, and L. Merino. “Decision-theoretic planning with person trajectory prediction for social navigation”. In: *Robot 2015: Second Iberian Robotics Conference: Advances in Robotics, Volume 2*. Springer. 2016, pp. 247–258.

[99] P. Trautman, J. Ma, R. M. Murray, and A. Krause. “Robot navigation in dense human crowds: the case for cooperation”. In: *2013 IEEE international conference on robotics and automation*. IEEE. 2013, pp. 2153–2160.

[100] L. Bruckschen, K. Bungert, N. Dengler, and M. Bennewitz. “Human-aware robot navigation by long-term movement prediction”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2020, pp. 11032–11037.

[101] M. Kamezaki, A. Kobayashi, R. Kono, M. Hirayama, and S. Sugano. “Dynamic Waypoint Navigation: Model-Based Adaptive Trajectory Planner for Human-Symbiotic Mobile Robots”. In: *IEEE Access* 10 (2022), pp. 81546–81555.

[102] K. Li, Y. Xu, J. Wang, and M. Q.-H. Meng. “SARL: Deep reinforcement learning based human-aware navigation for mobile robot in indoor environments”. In: *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE. 2019, pp. 688–694.

[103] A. K. Pandey and R. Alami. “A framework towards a socially aware mobile robot motion in human-centered dynamic environment”. In: *2010 IEEE/RSJ international Conference on Intelligent Robots and systems*. IEEE. 2010, pp. 5855–5860.

[104] Y. Chen, C. Liu, B. E. Shi, and M. Liu. “Robot navigation in crowds by graph convolutional networks with attention learned from human gaze”. In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 2754–2761.

[105] N. Fernández Coleto, E. Ruiz Ramírez, F. Haarslev, and L. Bodenhagen. “Towards socially acceptable, human-aware robot navigation”. In: Springer. 2019, pp. 578–587.

[106] K. Zhao, L. Zhou, Z. Hu, S. Cheng, A. Shi, Y. Sun, and J. Liu. “Human-Aware Robot Navigation Based on Asymmetric Gaussian Model”. In: *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE. 2021, pp. 322–327.

[107] Y. Chen and Y. Lou. “A unified multiple-motion-mode framework for socially compliant navigation in dense crowds”. In: *IEEE Transactions on Automation Science and Engineering* 19.4 (2021), pp. 3536–3548.

[108] A. Bera, T. Randhavane, R. Prinja, and D. Manocha. “Sociosense: Robot navigation amongst pedestrians with social and psychological constraints”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 7018–7025.

[109] J. Chen, X. Chen, and S. Liu. “Trajectory Planning of Autonomous Mobile Robot using Model Predictive Control in Human-Robot Shared Workspace”. In: *2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI)*. IEEE. 2023, pp. 462–467.

[110] M. Mayyas, S. P. Vadlamudi, and M. A. Syed. “Fenceless obstacle avoidance method for efficient and safe human–robot collaboration in a shared work space”. In: *International Journal of Advanced Robotic Systems* 17.5 (2020), p. 1729881420959018.

[111] S. B. Liu, H. Roehm, C. Heinzemann, I. Lütkebohle, J. Oehlerking, and M. Althoff. “Provably safe motion of mobile robots in human environments”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 1351–1357.

[112] J. Guzzi, A. Giusti, L. M. Gambardella, G. Theraulaz, and G. A. Di Caro. “Human-friendly robot navigation in dynamic environments”. In: *2013 IEEE international conference on robotics and automation*. IEEE. 2013, pp. 423–430.

[113] H. Kivrak, F. Cakmak, H. Kose, and S. Yavuz. “Waypoint based path planner for socially aware robot navigation”. In: *Cluster Computing* 25.3 (2022), pp. 1665–1675.

[114] P. Ratsamee, Y. Mae, K. Kamiyama, M. Horade, M. Kojima, and T. Arai. “Social interactive robot navigation based on human intention analysis from face orientation and human path prediction”. In: *ROBOMECH Journal* 2.1 (2015). ISSN: 21974225. DOI: 10.1186/s40648-015-0033-z.

[115] M. Galvan, E. Repiso, and A. Sanfeliu. “Robot Navigation to Approach People Using-Spline Path Planning and Extended Social Force Model”. In: *Robot 2019: Fourth Iberian Robotics Conference: Advances in Robotics, Volume 2*. Springer. 2019, pp. 15–27.

[116] G. Ferrer, A. Garrell, and A. Sanfeliu. “Social-aware robot navigation in urban environments”. In: *2013 European Conference on Mobile Robots*. IEEE. 2013, pp. 331–336.

[117] G. Ferrer, A. Garrell, and A. Sanfeliu. “Robot companion: A social-force based approach with human awareness-navigation in crowded environments”. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2013, pp. 1688–1694.

6

[118] P. Ratsamee, Y. Mae, K. Ohara, M. Kojima, and T. Arai. "Social navigation model based on human intention analysis using face orientation". In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2013, pp. 1682–1687.

[119] P. Regier, I. Shareef, and M. Bennewitz. "Improving navigation with the social force model by learning a neural network controller in pedestrian crowds". In: *2019 European Conference on Mobile Robots (ECMR)*. IEEE. 2019, pp. 1–6.

[120] M. Kamezaki, Y. Tsuburaya, T. Kanada, M. Hirayama, and S. Sugano. "Reactive, proactive, and inducible proximal crowd robot navigation method based on inducible social force model". In: *IEEE Robotics and Automation Letters* 7.2 (2022), pp. 3922–3929.

[121] G. Ferrer and A. Sanfeliu. "Proactive kinodynamic planning using the extended social force model and human motion prediction in urban environments". In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2014, pp. 1730–1735.

[122] M. Fuad, T. Agustinah, and D. Purwanto. "Modified headed social force model based on hybrid velocity obstacles for mobile robot to avoid disturbed groups of pedestrians". In: *International Journal of Intelligent Engineering and Systems* 14.3 (2021), pp. 222–241.

[123] Ó. Gil, A. Garrell, and A. Sanfeliu. "Social robot navigation tasks: Combining machine learning techniques and social force model". In: *Sensors* 21.21 (2021), p. 7087.

[124] E. Repiso, A. Garrell, and A. Sanfeliu. "Adaptive side-by-side social robot navigation to approach and interact with people". In: *International Journal of Social Robotics* 12 (2020), pp. 909–930.

[125] S. Oli, B. L'Esperance, and K. Gupta. "Human Motion Behaviour Aware Planner (HMBAP) for path planning in dynamic human environments". In: *2013 16th international conference on advanced robotics (ICAR)*. IEEE. 2013, pp. 1–7.

[126] S. Thompson, T. Horiuchi, and S. Kagami. "A probabilistic model of human motion and navigation intent for mobile robot path planning". In: *2009 4th International Conference on Autonomous Robots and Agents*. IEEE. 2009, pp. 663–668.

[127] A. Malviya and R. Kala. "Social robot motion planning using contextual distances observed from 3D human motion tracking". In: *Expert Systems with Applications* 184 (2021), p. 115515.

[128] C. Park, J. Ondřej, M. Gilbert, K. Freeman, and C. O'Sullivan. "Hi robot: Human intention-aware robot planning for safe and efficient navigation in crowds". In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2016, pp. 3320–3326.

[129] M. Pfeiffer, U. Schwesinger, H. Sommer, E. Galceran, and R. Siegwart. “Predicting actions to act predictably: Cooperative partial motion planning with maximum entropy models”. In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2016, pp. 2096–2101.

[130] S. Kodagoda, S. Sehestedt, and G. Dissanayake. “Socially aware path planning for mobile robots”. In: *Robotica* 34.3 (2016), pp. 513–526.

[131] J. Cai, A. Du, X. Liang, and S. Li. “Prediction-Based Path Planning for Safe and Efficient Human–Robot Collaboration in Construction via Deep Reinforcement Learning”. In: *Journal of Computing in Civil Engineering* 37.1 (2023), p. 04022046.

[132] I. Kostavelis, D. Giakoumis, S. Malassiotis, and D. Tzovaras. “Human aware robot navigation in semantically annotated domestic environments”. In: *Universal Access in Human-Computer Interaction. Interaction Techniques and Environments: 10th International Conference, UAHCI 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings, Part II 10*. Springer. 2016, pp. 414–423.

[133] T. Smith, Y. Chen, N. Hewitt, B. Hu, and Y. Gu. “Socially aware robot obstacle avoidance considering human intention and preferences”. In: *International journal of social robotics* (2021), pp. 1–18.

[134] A. Nagariya, B. Gopalakrishnan, A. K. Singh, K. Gupta, and K. M. Krishna. “Mobile robot navigation amidst humans with intents and uncertainties: A time scaled collision cone approach”. In: *2015 54th IEEE Conference on Decision and Control (CDC)*. 2015, pp. 2773–2779.

[135] Y. Lu, X. Ruan, and J. Huang. “Deep reinforcement learning based on social spatial–temporal graph convolution network for crowd navigation”. In: *Machines* 10.8 (2022), p. 703.

[136] V. K. Narayanan, T. Miyashita, and N. Hagita. “Formalizing a transient-goal driven approach for pedestrian-aware robot navigation”. In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2018, pp. 862–867.

[137] V. V. Unhelkar, C. Pérez-D’Arpino, L. Stirling, and J. A. Shah. “Human-robot co-navigation using anticipatory indicators of human walking motion”. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2015, pp. 6183–6190.

[138] N. Nishino, R. Tsugita, D. Chugo, S. Muramatsu, S. Yokota, and H. Hashimoto. “Mobile robot considering a pedestrian walking model based on the environmental characteristics”. In: *2017 IEEE/SICE International Symposium on System Integration (SII)*. IEEE. 2017, pp. 716–721.

[139] N. Sakata, Y. Kinoshita, and Y. Kato. “Predicting a pedestrian trajectory using seq2seq for mobile robot navigation”. In: *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*. IEEE. 2018, pp. 4300–4305.

[140] L. Sun, Z. Yan, S. M. Mellado, M. Hanheide, and T. Duckett. "3DOF pedestrian trajectory prediction learned from long-term autonomous mobile robot deployment data". In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 5942–5948.

[141] S. Schaefer, K. Leung, B. Ivanovic, and M. Pavone. "Leveraging neural network gradients within trajectory optimization for proactive human-robot interactions". In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2021, pp. 9673–9679.

[142] H. Nishimura, B. Ivanovic, A. Gaidon, M. Pavone, and M. Schwager. "Risk-sensitive sequential action control with multi-modal human trajectory forecasting for safe crowd-robot interaction". In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2020, pp. 11205–11212.

[143] Z. Chen, C. Song, Y. Yang, B. Zhao, Y. Hu, S. Liu, and J. Zhang. "Robot navigation based on human trajectory prediction and multiple travel modes". In: *Applied Sciences* 8.11 (2018), p. 2205.

[144] T. Salzmann, H.-T. L. Chiang, M. Ryll, D. Sadigh, C. Parada, and A. Bewley. "Robots That Can See: Leveraging Human Pose for Trajectory Prediction". In: *IEEE Robotics and Automation Letters* (2023).

[145] Y. Sato, Y. Sasaki, and H. Takemura. "STP4: spatio temporal path planning based on pedestrian trajectory prediction in dense crowds". In: *PeerJ Computer Science* 9 (2023), e1641.

[146] S.-H. Wang, Y.-H. Wu, and T.-H. S. Li. "Deep Reinforcement Learning with Pedestrian Trajectory Prediction Model for Service Robot Navigation in Crowded Environments". In: *2023 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*. IEEE. 2023, pp. 1–6.

[147] S. Liu, P. Chang, Z. Huang, N. Chakraborty, K. Hong, W. Liang, D. L. McPherson, J. Geng, and K. Driggs-Campbell. "Intention aware robot crowd navigation with attention-based interaction graph". In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2023, pp. 12015–12021.

[148] J. Long, J. Mei, and G. Ma. "Egocentric Two-Frame Pedestrian Trajectory Prediction Algorithm Based on a Panoramic Camera". In: *IEEE Transactions on Instrumentation and Measurement* 72 (2022), pp. 1–13.

[149] C. Weinrich, M. Volkhardt, E. Einhorn, and H.-M. Gross. "Prediction of human collision avoidance behavior by lifelong learning for socially compliant robot navigation". In: *2013 IEEE International Conference on Robotics and Automation*. IEEE. 2013, pp. 376–381.

[150] Y. Kobayashi, T. Sugimoto, K. Tanaka, Y. Shimomura, F. J. Arjonilla Garcia, C. H. Kim, H. Yabushita, and T. Toda. "Robot navigation based on predicting of human interaction and its reproducible evaluation in a densely crowded environment". In: *International Journal of Social Robotics* (2022), pp. 1–15.

[151] V. Narayanan, B. M. Manoghar, R. P. RV, and A. Bera. “Ewarenet: Emotion-aware pedestrian intent prediction and adaptive spatial profile fusion for social robot navigation”. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2023, pp. 7569–7575.

[152] C. Lichtenthäler, T. Lorenzy, and A. Kirsch. “Influence of legibility on perceived safety in a virtual human-robot path crossing task”. In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2012, pp. 676–681.

[153] T. Kruse, P. Basili, S. Glasauer, and A. Kirsch. “Legible robot navigation in the proximity of moving humans”. In: *2012 IEEE workshop on advanced robotics and its social impacts (ARSO)*. IEEE. 2012, pp. 83–88.

[154] F. Papenmeier, M. Uhrig, and A. Kirsch. “Human understanding of robot motion: the role of velocity and orientation”. In: *International Journal of Social Robotics* 11 (2019), pp. 75–88.

[155] O. Palinko, E. R. Ramírez, W. K. Juel, N. Krüger, and L. Bodenhagen. “Intention Indication for Human Aware Robot Navigation”. In: *VISIGRAPP (2: HUCAPP)*. 2020, pp. 64–74.

[156] Y. Che, A. M. Okamura, and D. Sadigh. “Efficient and trustworthy social navigation via explicit and implicit robot–human communication”. In: *IEEE Transactions on Robotics* 36.3 (2020), pp. 692–707.

[157] A. D. May, C. Dondrup, and M. Hanheide. “Show me your moves! Conveying navigation intention of a mobile robot to humans”. In: *2015 European Conference on Mobile Robots (ECMR)*. IEEE. 2015, pp. 1–6.

[158] B. Lewandowski, T. Wengefeld, S. Muller, M. Jenny, S. Glende, C. Schroter, A. Bley, and H. M. Gross. “Socially Compliant Human-Robot Interaction for Autonomous Scanning Tasks in Supermarket Environments”. In: *29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2020* 1 (2020), pp. 363–370. DOI: 10.1109/RO-MAN47096.2020.9223568.

[159] Y. Che, C. T. Sun, and A. M. Okamura. “Avoiding human-robot collisions using haptic communication”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 5828–5834.

[160] D. Carton, W. Olszowy, D. Wollherr, and M. Buss. “Socio-contextual constraints for human approach with a mobile robot”. In: *International Journal of Social Robotics* 9 (2017), pp. 309–327.

[161] S. S. Kannan, A. Lee, and B.-C. Min. “External Human-Machine Interface on Delivery Robots: Expression of Navigation Intent of the Robot”. In: *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE. 2021, pp. 1305–1312.

[162] H.-Y. Chen, P.-H. Huang, and L.-C. Fu. “Social crowd navigation of a mobile robot based on human trajectory prediction and hybrid sensing”. In: *Autonomous Robots* (2023), pp. 1–13.

[163] A. Bera, T. Randhavane, and D. Manocha. “Improving Socially-aware Multi-channel Human Emotion Prediction for Robot Navigation”. In: *CVPR Workshops*. 2019, pp. 21–27.

[164] A. Bera, T. Randhavane, and D. Manocha. “Modelling multi-channel emotions using facial expression and trajectory cues for improving socially-aware robot navigation”. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* 2019-June (2019), pp. 257–266. ISSN: 21607516. DOI: 10.1109/CVPRW.2019.00035.

[165] A. Bera, T. Randhavane, R. Prinja, K. Kapsaskis, A. Wang, K. Gray, and D. Manocha. “How are you feeling? multimodal emotion learning for socially-assistive robot navigation”. In: *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE. 2020, pp. 644–651.

[166] V. Narayanan, B. M. Manoghar, V. S. Dorbala, D. Manocha, and A. Bera. “ProxEmo: Gait-based Emotion Learning and Multi-view Proxemic Fusion for Socially-Aware Robot Navigation”. In: (2020).

[167] A. Cosgun, E. A. Sisbot, and H. I. Christensen. “Anticipatory robot path planning in human environments”. In: *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE. 2016, pp. 562–569.

[168] M. Daza, D. Barrios-Aranibar, J. Diaz-Amado, Y. Cardinale, and J. Vilasboas. “An approach of social navigation based on proxemics for crowded environments of humans and robots”. In: *Micromachines* 12.2 (2021), p. 193.

[169] A. Vega, L. J. Manso, D. G. 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”. In: *Pattern Recognition Letters* 118 (2019), pp. 72–84.

[170] R. M. d. Sousa, D. Barrios-Aranibar, J. Diaz-Amado, R. E. Patiño-Escarcina, and R. M. P. Trindade. “A New Approach for Including Social Conventions into Social Robots Navigation by Using Polygonal Triangulation and Group Asymmetric Gaussian Functions”. In: *Sensors* 22.12 (2022), p. 4602.

[171] X. T. Truong, Y. S. Ou, and T.-D. Ngo. “Towards culturally aware robot navigation”. In: *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE. 2016, pp. 63–69.

[172] A. J. Sathyamoorthy, U. Patel, M. Paul, N. K. S. Kumar, Y. Savle, and D. Manocha. “CoMet: Modeling group cohesion for socially compliant robot navigation in crowded scenes”. In: *IEEE Robotics and Automation Letters* 7.2 (2021), pp. 1008–1015.

[173] A. Kollakidou, L. Naik, O. Palinko, and L. Bodenhagen. “Enabling Robots to Adhere to Social Norms by Detecting F-Formations”. In: *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE. 2021, pp. 110–116.

[174] N. A. Pham, L. A. Nguyen, and X. T. Truong. "Socially aware robot navigation framework: Social activities recognition using deep learning techniques". In: *2021 8th NAFOSTED Conference on Information and Computer Science (NICS)*. IEEE. 2021, pp. 381–385.

[175] Y. Nakauchi and R. Simmons. "A service robot for peopled environments-Finding lined people by stereo vision and standing in line". In: *2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Proceedings (Cat. No. 01TH8556)*. Vol. 2. IEEE. 2001, pp. 1093–1098.

[176] K. Charalampous, I. Kostavelis, and A. Gasteratos. "Robot navigation in large-scale social maps: An action recognition approach". In: *Expert Systems with Applications* 66 (2016), pp. 261–273.

[177] A. C. Ramirez-Hernandez, J. A. Rivera-Bautista, A. Marin-Hernandez, and V. A. Garcia-Vega. "Detection and interpretation of human walking gestures for human-robot interaction". In: *2009 Eighth Mexican International Conference on Artificial Intelligence*. IEEE. 2009, pp. 41–46.

[178] J. Vroon, G. Englebienne, and V. Evers. "Detecting perceived appropriateness of a robot's social positioning behavior from non-verbal cues". In: *Proceedings - 2019 IEEE 1st International Conference on Cognitive Machine Intelligence, CogMI 2019* (2019), pp. 216–225. DOI: 10.1109/CogMI48466.2019.00039.

[179] F. Fang, M. Shi, K. Qian, B. Zhou, and Y. Gan. "A human-aware navigation method for social robot based on multi-layer cost map". In: *International Journal of Intelligent Robotics and Applications* 4 (2020), pp. 308–318.

[180] X. Zhou, X. Ye, K. Zhang, and S. Yu. "Deep Reinforcement Learning for Group-Aware Robot Navigation in Crowds". In: *International Conference on Knowledge Science, Engineering and Management*. Springer. 2023, pp. 25–34.

[181] S. F. Warta, O. B. Newton, J. Song, A. Best, and S. M. 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. 1. SAGE Publications Sage CA: Los Angeles, CA. 2018, pp. 1128–1132.

[182] T. J. Wiltshire, E. J. Lobato, A. V. Wedell, W. Huang, B. Axelrod, and S. M. Fiore. "Effects of robot gaze and proxemic behavior on perceived social presence during a hallway navigation scenario". In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 57. 1. SAGE Publications Sage CA: Los Angeles, CA. 2013, pp. 1273–1277.

[183] L. J. Manso, R. R. Jorvekar, D. R. Faria, P. Bustos, and P. Bachiller. "Graph neural networks for human-aware social navigation". In: *Advances in Physical Agents II: Proceedings of the 21st International Workshop of Physical Agents (WAF 2020), November 19-20, 2020, Alcalá de Henares, Madrid, Spain*. Springer. 2021, pp. 167–179.

[184] A. Sorrentino, O. Khalid, L. Coviello, F. Cavallo, and L. Fiorini. “Modeling human-like robot personalities as a key to foster socially aware navigation”. In: *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE. 2021, pp. 95–101.

[185] K. He, W. P. Chan, A. Cosgun, A. Joy, and E. A. Croft. “Robot Gaze During Autonomous Navigation and Its Effect on Social Presence”. In: *International Journal of Social Robotics* (2023), pp. 1–19.

[186] A. T. Rifqi, B. S. B. Dewantara, D. Pramadihanto, and B. S. Marta. “Fuzzy Social Force Model for Healthcare Robot Navigation and Obstacle Avoidance”. In: *2021 International Electronics Symposium (IES)*. IEEE. 2021, pp. 445–450.

[187] T. Kirks, J. Jost, J. Finke, and S. Hoose. “Modelling Proxemics for Human-Technology-Interaction in Decentralized Social-Robot-Systems”. In: *Intelligent Human Systems Integration 2020: Proceedings of the 3rd International Conference on Intelligent Human Systems Integration (IHSI 2020): Integrating People and Intelligent Systems, February 19-21, 2020, Modena, Italy*. Springer. 2020, pp. 153–158.

[188] J. Montesdeoca, J. M. Toibero, J. Jordan, A. Zell, and R. Carelli. “Person-Following Controller with Socially Acceptable Robot Motion”. In: *Robotics and Autonomous Systems* 153 (2022), p. 104075.

[189] C. Wang, Y. Li, S. S. Ge, and T. H. Lee. “Adaptive control for robot navigation in human environments based on social force model”. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2016, pp. 5690–5695.

[190] M. Svenstrup, S. T. Hansen, H. J. Andersen, and T. Bak. “Adaptive human-aware robot navigation in close proximity to humans”. In: *International Journal of Advanced Robotic Systems* 8.2 (2011), p. 21.

[191] X.-T. Truong and T.-D. Ngo. “Dynamic social zone based mobile robot navigation for human comfortable safety in social environments”. In: *International Journal of Social Robotics* 8.5 (2016), pp. 663–684.

[192] M. A. Kenk, M. Hassaballah, and J.-F. Brethé. “Human-aware Robot Navigation in Logistics Warehouses.” In: *ICINCO (2)*. 2019, pp. 371–378.

[193] K. Tatarian, R. Stower, D. Rudaz, M. Chamoux, A. Kappas, and M. Chetouani. “How does modality matter? investigating the synthesis and effects of multi-modal robot behavior on social intelligence”. In: *International Journal of Social Robotics* 14.4 (2022), pp. 893–911.

[194] S. Forer, S. B. Banisetty, L. Yliniemi, M. Nicolescu, and D. Feil-Seifer. “Socially-aware navigation using non-linear multi-objective optimization”. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2018, pp. 1–9.

[195] S. B. Banisetty, S. Forer, L. Yliniemi, M. Nicolescu, and D. Feil-Seifer. “Socially aware navigation: A non-linear multi-objective optimization approach”. In: *ACM Transactions on Interactive Intelligent Systems (TiiS)* 11.2 (2021), pp. 1–26.

[196] M. Liu, L. Zhou, Y. Sun, and J. Liu. “Robot Comfort Following based on Extended Social Force Model in Dynamic Environment”. In: *2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE. 2019, pp. 30–35.

[197] M. Svenstrup, S. Tranberg, H. J. Andersen, and T. Bak. “Pose estimation and adaptive robot behaviour for human-robot interaction”. In: *Proceedings - IEEE International Conference on Robotics and Automation* (2009), pp. 3571–3576. ISSN: 10504729. DOI: 10.1109/ROBOT.2009.5152690.

[198] J. C. Montesdeoca, M. Toibero, and R. Carelli. “Person-following based on social navigation into the sensorized environments”. In: *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE. 2017, pp. 799–803.

[199] W. Zhu and M. Hayashibe. “Autonomous Navigation System in Pedestrian Scenarios Using a Dreamer-Based Motion Planner”. In: *IEEE Robotics and Automation Letters* (2023).

[200] R. Jiang, S. S. Ge, N. T. Tangirala, and T. H. Lee. “Interactive navigation of mobile robots based on human’s emotion”. In: *International Conference on Social Robotics*. Springer. 2016, pp. 243–252.

[201] Y. Nakauchi and R. Simmons. “Social behavioral robot that stands in line”. In: *IEEE SMC’99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028)*. Vol. 2. IEEE. 1999, pp. 993–998.

[202] S. B. Vatan, K. Bektaş, and H. I. Bozma. “Social APF-RL: Safe Mapless Navigation in Unknown & Human-Populated Environments”. In: *2023 European Conference on Mobile Robots (ECMR)*. IEEE. 2023, pp. 1–6.

[203] E. Pacchierotti, H. I. Christensen, and P. Jensfelt. “Embodied social interaction for service robots in hallway environments”. In: *Springer Tracts in Advanced Robotics* 25.2000 (2006), pp. 293–304. ISSN: 1610742X. DOI: 10.1007/978-3-540-33453-8_25.

[204] A. Bellarbi, A.-i. Mouaddib, N. Achour, and N. Ouadah. “A new approach for social navigation and interaction using a dynamic proxemia modeling”. In: *Evolutionary Intelligence* (2022), pp. 1–27.

[205] L. V. Calderita, A. Vega, P. Bustos, and P. Núñez. “A new human-aware robot navigation framework based on time-dependent social interaction spaces: An application to assistive robots in caregiving centers”. In: *Robotics and Autonomous Systems* 145 (2021), p. 103873.

[206] A. Vega, R. Cintas, L. J. Manso, P. Bustos, and P. Núñez. “Socially-accepted path planning for robot navigation based on social interaction spaces”. In: *Robot 2019: Fourth Iberian Robotics Conference: Advances in Robotics, Volume 2*. Springer. 2020, pp. 644–655.

6

[207] A. Vega-Magro, L. Manso, P. Bustos, P. Nunez, and D. G. Macharet. “Socially acceptable robot navigation over groups of people”. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2017, pp. 1182–1187.

[208] A. Vega-Magro, L. V. Calderita, P. Bustos, and P. Núñez. “Human-aware robot navigation based on time-dependent social interaction spaces: A use case for assistive robotics”. In: *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. IEEE. 2020, pp. 140–145.

[209] L. V. Calderita, A. Vega, P. Bustos, and P. Núñez. “Social Robot Navigation adapted to Time-dependent Affordance Spaces: A Use Case for Caregiving Centers”. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2020, pp. 944–949.

[210] M. Garcia-Salguero, J. Monroy, A. Solano, and J. Gonzalez-Jimenez. “Socially acceptable approach to humans by a mobile robot”. In: *Proceedings of the 2nd International Conference on Applications of Intelligent Systems*. 2019, pp. 1–7.

[211] A. Mora, A. Prados, A. Mendez, R. Barber, and S. Garrido. “Sensor Fusion for Social Navigation on a Mobile Robot Based on Fast Marching Square and Gaussian Mixture Model”. In: *Sensors* 22.22 (2022), p. 8728.

[212] X.-T. Truong and T.-D. Ngo. ““To approach humans?”: A unified framework for approaching pose prediction and socially aware robot navigation”. In: *IEEE Transactions on Cognitive and Developmental Systems* 10.3 (2017), pp. 557–572.

[213] M. Kollmitz, K. Hsiao, J. Gaa, and W. Burgard. “Time dependent planning on a layered social cost map for human-aware robot navigation”. In: *2015 European Conference on Mobile Robots (ECMR)*. IEEE. 2015, pp. 1–6.

[214] Y. Nakauchi and R. Simmons. “A social robot that stands in line”. In: *Autonomous Robots* 12.3 (2002), pp. 313–324. ISSN: 09295593. DOI: 10.1023/A:1015273816637.

[215] E. Torta, R. H. Cuijpers, and J. F. Juola. “Design of a parametric model of personal space for robotic social navigation”. In: *International Journal of Social Robotics* 5 (2013), pp. 357–365.

[216] Z. Talebpour, I. Navarro, and A. Martinoli. “On-board human-aware navigation for indoor resource-constrained robots: A case-study with the ranger”. In: *2015 IEEE/SICE International Symposium on System Integration (SII)*. IEEE. 2015, pp. 63–68.

[217] M. M. Neggers, R. H. Cuijpers, P. A. Ruijten, and W. A. IJsselsteijn. “Determining Shape and Size of Personal Space of a Human when Passed by a Robot”. In: *International Journal of Social Robotics* 14.2 (2022), pp. 561–572.

[218] S. Silva, D. Paillacho, N. Verdezoto, and J. D. Hernández. “Towards online socially acceptable robot navigation”. In: *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*. IEEE. 2022, pp. 707–714.

[219] A. Mateus, P. Miraldo, P. U. Lima, and J. Sequeira. "Human-aware navigation using external omnidirectional cameras". In: *Robot 2015: Second Iberian Robotics Conference: Advances in Robotics, Volume 1*. Springer. 2015, pp. 283–295.

[220] H. Q. T. Ngo, V. N. Le, V. D. N. Thien, T. P. Nguyen, and H. Nguyen. "Develop the socially human-aware navigation system using dynamic window approach and optimize cost function for autonomous medical robot". In: *Advances in Mechanical Engineering* 12.12 (2020), p. 1687814020979430.

[221] I. Nishitani, T. Matsumura, M. Ozawa, A. Yorozu, and M. Takahashi. "Human-centered X–Y–T space path planning for mobile robot in dynamic environments". In: *Robotics and Autonomous Systems* 66 (2015), pp. 18–26.

[222] K.-C. Huang, J.-Y. Li, and L.-C. Fu. "Human-oriented navigation for service providing in home environment". In: *Proceedings of SICE Annual Conference 2010*. IEEE. 2010, pp. 1892–1897.

[223] S. F. Chik, C. F. Yeong, L. M. Su, T. Y. Lim, F. Duan, and J. H. Chin. "Neural-network based adaptive proxemics-costmap for human-aware autonomous robot navigation". In: *International Journal of Integrated Engineering* 11.4 (2019).

[224] A. Toyoshima, N. Nishino, D. Chugo, S. Muramatsu, S. Yokota, and H. Hashimoto. "Autonomous Mobile Robot Navigation: Consideration of the Pedestrian's Dynamic Personal Space". In: *2018 IEEE 27th International Symposium on Industrial Electronics (ISIE)*. IEEE. 2018, pp. 1094–1099.

[225] W. Chi, H. Kono, Y. Tamura, A. Yamashita, H. Asama, and M. Q.-H. Meng. "A human-friendly robot navigation algorithm using the risk-RRT approach". In: *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE. 2016, pp. 227–232.

[226] T. D. Ngo, X.-T. Truong, *et al.* "Socially aware robot navigation framework: Where and how to approach people in dynamic social environments". In: *IEEE Transactions on Automation Science and Engineering* 20.2 (2022), pp. 1322–1336.

[227] S. F. Chik, C. F. Yeong, E. L. M. Su, T. Y. Lim, F. Duan, J. T. C. Tan, P. H. Tan, and P. J. H. Chin. "Gaussian pedestrian proxemics model with social force for service robot navigation in dynamic environment". In: *Asian Simulation Conference*. Springer. 2017, pp. 61–73.

[228] D. N. Do, V. X. Nguyen, X.-T. Truong, *et al.* "An integrated avoiding and approaching human systems for mobile service robots in dynamic social environments". In: *Vietnam Journal of Science and Technology* 60.1 (2022), pp. 114–126.

[229] C. A. My, X. T. Truong, *et al.* "Toward socially aware trajectory planning system for autonomous mobile robots in complex environments". In: *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*. IEEE. 2020, pp. 90–95.

[230] X.-T. Truong, V. N. Yoong, and T.-D. Ngo. “Dynamic social zone for human safety in human-robot shared workspaces”. In: *2014 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. IEEE. 2014, pp. 391–396.

[231] O. Palinko, E. R. Ramirez, N. Krüger, and L. Bodenhagen. “Intention Understanding for Human-Aware Mobile Robots: Comparing Cues and the Effect of Demographics”. In: *International Joint Conference on Computer Vision, Imaging and Computer Graphics*. Springer. 2020, pp. 83–102.

[232] V. V. Unhelkar, H. C. Siu, and J. A. Shah. “Comparative performance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks”. In: *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2014, pp. 82–89.

[233] A. M. Abrams, P. S. Dautzenberg, C. Jakobowsky, S. Ladwig, and A. M. Rosenthal-von der Pütten. “A Theoretical and Empirical Reflection on Technology Acceptance Models for Autonomous Delivery Robots”. In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 2021, pp. 272–280.

[234] A. A. Hayat, R. Parween, M. R. Elara, K. Parsuraman, and P. S. Kandasamy. “Panthera: Design of a reconfigurable pavement sweeping robot”. In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 2019, pp. 7346–7352.

[235] J. Guerreiro, D. Sato, S. Asakawa, H. Dong, K. M. Kitani, and C. Asakawa. “Cabot: Designing and evaluating an autonomous navigation robot for blind people”. In: *The 21st International ACM SIGACCESS conference on computers and accessibility*. 2019, pp. 68–82.

[236] S. Wang and H. I. Christensen. “Tritonbot: First lessons learned from deployment of a long-term autonomy tour guide robot”. In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2018, pp. 158–165.

[237] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita. “How to approach humans?-strategies for social robots to initiate interaction”. In: *Human-Robot Interaction (HRI), 2009 4th ACM/IEEE International Conference on*. IEEE. 2009, pp. 109–116.

[238] S. Satake, K. Hayashi, K. Nakatani, and T. Kanda. “Field trial of an information-providing robot in a shopping mall”. In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2015, pp. 1832–1839.

[239] C. Bennett, E. Ackerman, B. Fan, J. Bigham, P. Carrington, and S. Fox. “Accessibility and The Crowded Sidewalk: Micromobility’s Impact on Public Space”. In: *Designing Interactive Systems Conference 2021*. 2021, pp. 365–380.

[240] D. Brščić, H. Kidokoro, Y. Suehiro, and T. Kanda. “Escaping from children’s abuse of social robots”. In: *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction*. 2015, pp. 59–66.

[241] P. Salvini, G. Ciaravella, W. Yu, G. Ferri, A. Manzi, B. Mazzolai, C. Laschi, S.-R. Oh, and P. Dario. “How safe are service robots in urban environments? Bullying a robot”. In: *19th international symposium in robot and human interactive communication*. IEEE. 2010, pp. 1–7.

[242] M. Bruijnes, R. van Delden, and J. H. Vroon. “Telepresence Robots in the Wide Wild World”. In: *3rd Workshop on Ethical Encounters in HCI: Research in Sensitive Settings*. 2017.

[243] G. Nichols. *San Francisco bans delivery robots in most of the city*. 2017. URL: https://www.zdnet.com/article/san-francisco-bans-delivery-robots-in-most-of-the-city/.

[244] T. Schulz, P. Holthaus, F. Amirabdollahian, and K. L. Koay. “Humans’ perception of a robot moving using a slow in and slow out velocity profile”. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2019, pp. 594–595.

[245] L. Moshkina, S. Trickett, and J. G. Trafton. “Social engagement in public places: a tale of one robot”. In: *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. 2014, pp. 382–389.

[246] 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 International Conference on Human-Robot Interaction (HRI)*. IEEE. 2019, pp. 421–430.

[247] E. Senft, S. Satake, and T. Kanda. “Would You Mind Me if I Pass by You? Socially-Appropriate Behaviour for an Omni-based Social Robot in Narrow Environment”. In: *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 2020, pp. 539–547.

[248] C. I. Mavrogiannis, W. B. Thomason, and R. A. Knepper. “Social momentum: A framework for legible navigation in dynamic multi-agent environments”. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 2018, pp. 361–369.

[249] A. Rosenthal-von der Pütten, D. Sirkin, A. Abrams, and L. Platte. “The Forgotten in HRI: Incidental Encounters with Robots in Public Spaces”. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 2020, pp. 656–657.

[250] T. Rosenbloom, D. Nemrodov, and A. B. Eliyahu. “Yielding behavior of Israeli drivers: interaction of age and sex”. In: *Perceptual and motor skills* 103.2 (2006), pp. 387–390.

[251] X. Kong, S. Das, Y. Zhang, and X. Xiao. “Lessons learned from pedestrian-driver communication and yielding patterns”. In: *Transportation research part F: traffic psychology and behaviour* 79 (2021), pp. 35–48.

[252] X. Zhuang and C. Wu. “Pedestrian gestures increase driver yielding at uncontrolled mid-block road crossings”. In: *Accident Analysis & Prevention* 70 (2014), pp. 235–244.

[253] J. Vroon, Z. Rusák, and G. Kortuem. “Context-Confrontation: Elicitation and Exploration of Conflicts for Delivery Robots on Sidewalks”. In: *29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN 2020). First international workshop on Designerly HRI Knowledge.* 2020.

[254] H. Murakami, C. Feliciani, Y. Nishiyama, and K. Nishinari. “Mutual anticipation can contribute to self-organization in human crowds”. In: *Science Advances* 7.12 (2021), eabe7758.

[255] A. Corbetta, J. A. Meeusen, C.-m. Lee, R. Benzi, and F. Toschi. “Physics-based modeling and data representation of pairwise interactions among pedestrians”. In: *Physical Review E* 98.6 (2018), p. 062310.

[256] E. T. Hall. *The Hidden Dimension.* Anchor Books New York, 1966.

[257] J. R. Aiello and D. E. Thompson. “When compensation fails: Mediating effects of sex and locus of control at extended interaction distances”. In: *Basic and Applied Social Psychology* 1.1 (1980), pp. 65–82.

[258] A. Kendon. *Conducting interaction: Patterns of behavior in focused encounters.* Vol. 7. CUP Archive, 1990.

[259] R. Möller, A. Furnari, S. Battiato, A. Härmä, and G. M. Farinella. “A Survey on Human-aware Robot Navigation”. In: *arXiv preprint arXiv:2106.11650* (2021).

[260] J. Rios-Martinez, A. Spalanzani, and C. Laugier. “From proxemics theory to socially-aware navigation: A survey”. In: *International Journal of Social Robotics* 7.2 (2015), pp. 137–153.

[261] T. Kruse, A. Kirsch, E. A. Sisbot, and R. Alami. “Exploiting human cooperation in human-centered robot navigation”. In: *RO-MAN, 2010 IEEE.* IEEE. 2010, pp. 192–197.

[262] B. Leichtmann and V. Nitsch. “How much distance do humans keep toward robots? Literature review, meta-analysis, and theoretical considerations on personal space in human-robot interaction”. In: *Journal of environmental Psychology* 68 (2020), p. 101386.

[263] T. Kruse, A. Kirsch, H. Khambhaita, and R. Alami. “Evaluating directional cost models in navigation”. In: *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction.* ACM. 2014, pp. 350–357.

[264] F. Lindner. “How to Count Multiple Personal-Space Intrusions in Social Robot Navigation”. In: *Robophilosophy 2016.* IOS Press. 2016.

[265] C. Lichtenthäler and A. Kirsch. “Goal-predictability vs. trajectory-predictability: which legibility factor counts”. In: *Proceedings of the 2014 acm/ieee international conference on human-robot interaction.* ACM. 2014, pp. 228–229.

[266] R. Mead and M. J. Mataric. “Robots Have Needs Too: People Adapt Their Proxemic Preferences to Improve Autonomous Robot Recognition of Human Social Signals”. In: *New Frontiers in Human-Robot Interaction* (2015), p. 100.

[267] T. Q. Trinh, C. Schroeter, J. Kessler, and H.-M. Gross. "“Go ahead, please”: recognition and resolution of conflict situations in narrow passages for polite mobile robot navigation". In: *International Conference on Social Robotics*. Springer. 2015, pp. 643–653.

[268] M. Jung and P. Hinds. *Robots in the wild: A time for more robust theories of human-robot interaction*. 2018.

[269] K. Fischer, S. Yang, B. Mok, R. Maheshwari, D. Sirkin, and W. Ju. "Initiating interactions and negotiating approach: a robotic trash can in the field". In: *AAAI Symposium on Turn-taking and Coordination in Human-Machine Interaction*. AAAI Press. 2015, pp. 10–16.

[270] H. Hüttenrauch, K. Severinson Eklundh, A. Green, and E. A. Topp. "Investigating spatial relationships in human-robot interaction". In: *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. IEEE. 2006, pp. 5052–5059.

[271] H. Kidokoro, T. Kanda, D. Bršcic, and M. Shiomi. "Will I bother here? - A robot anticipating its influence on pedestrian walking comfort". In: *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 2013, pp. 259–266. DOI: 10.1109/HRI.2013.6483597.

[272] S. van Mierlo. "Field observations of reactions of incidentally copresent pedestrians to a seemingly autonomous sidewalk delivery vehicle: An exploratory study". MA thesis. the Netherlands: University of Utrecht, 2021.

[273] A. M. Abrams, L. Platte, and A. Rosenthal-von der Pütten. "Field observation: interactions between pedestrians and a delivery robot". In: (2020).

[274] P. Desmet. "Measuring emotion: Development and application of an instrument to measure emotional responses to products". In: *Funology*. Springer, 2003, pp. 111–123.

[275] P. A. Andersen, L. K. Guerrero, D. B. Buller, and P. F. Jorgensen. "An empirical comparison of three theories of nonverbal immediacy exchange". In: *Human Communication Research* 24.4 (1998), pp. 501–535.

[276] L. A. Hayduk. "Personal space: Where we now stand." In: *Psychological bulletin* 94.2 (1983), p. 293.

[277] M. Argyle and J. Dean. "Eye-contact, distance and affiliation". In: *Sociometry* (1965), pp. 289–304.

[278] R. Mead and M. J. Matarić. "Autonomous human–robot proxemics: socially aware navigation based on interaction potential". In: *Autonomous Robots* 41.5 (2017), pp. 1189–1201.

[279] D. C. May, K. J. Holler, C. L. Bethel, L. Strawderman, D. W. Carruth, and J. M. Usher. "Survey of factors for the prediction of human comfort with a non-anthropomorphic robot in public spaces". In: *International Journal of Social Robotics* 9 (2017), pp. 165–180.

6

[280] H. Khambhaita and R. Alami. "Viewing robot navigation in human environment as a cooperative activity". In: *Robotics Research: The 18th International Symposium ISRR*. Springer. 2020, pp. 285–300.

[281] A. Vinciarelli, M. Pantic, and H. Bourlard. "Social Signal Processing: Survey of an Emerging Domain". In: *Image and vision computing* 27.12 (2009). 10.1016/j.imavis.2008.11.007, pp. 1743–1759. ISSN: 0262-8856. DOI: 10.1016/j.imavis.2008.11.007.

[282] S. Saunderson and G. Nejat. "How robots influence humans: A survey of nonverbal communication in social human–robot interaction". In: *International Journal of Social Robotics* 11 (2019), pp. 575–608.

[283] M. Bretan, G. Hoffman, and G. Weinberg. "Emotionally expressive dynamic physical behaviors in robots". In: *International Journal of Human-Computer Studies* 78 (2015), pp. 1–16.

[284] N. Pérez-Higueras, R. Otero, F. Caballero, and L. Merino. "HuNavSim: A ROS 2 Human Navigation Simulator for Benchmarking Human-Aware Robot Navigation". In: *arXiv preprint arXiv:2305.01303* (2023).

[285] J. Vroon. "Responsive social positioning behaviour: for semi-autonomous telepresence robots". English. FP7-ICT-611153. PhD thesis. Netherlands: University of Twente, Sept. 2018. ISBN: 978-90-365-4619-5. DOI: 10.3990/1.9789036546195.

[286] I. Kastanis and M. Slater. "Reinforcement Learning Utilizes Proxemics: An Avatar Learns to Manipulate the Position of People in Immersive Virtual Reality". In: *ACM Trans. Appl. Percept.* 9.1 (Mar. 2012), 3:1–3:15. ISSN: 1544-3558. DOI: 10.1145/2134203.2134206. URL: http://doi.acm.org/10.1145/2134203.2134206.

[287] M. F. Jung. "Affective Grounding in Human-Robot Interaction". In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2017, pp. 263–273.

[288] J. Vroon, G. Englebienne, and V. Evers. "Detecting perceived appropriateness of a robot's social positioning behavior from non-verbal cues". In: *2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI)*. IEEE. 2019, pp. 216–225.

[289] C. Mavrogiannis, P. Alves-Oliveira, W. Thomason, and R. A. Knepper. "Social momentum: Design and evaluation of a framework for socially competent robot navigation". In: *ACM Transactions on Human-Robot Interaction (THRI)* 11.2 (2022), pp. 1–37.

[290] M. Kyrarini, F. Lygerakis, A. Rajavenkatanarayanan, C. Sevastopoulos, H. R. Nambiappan, K. K. Chaitanya, A. R. Babu, J. Mathew, and F. Makedon. "A survey of robots in healthcare". In: *Technologies* 9.1 (2021), p. 8.

[291] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh. "Core challenges of social robot navigation: A survey". In: *ACM Transactions on Human-Robot Interaction* 12.3 (2023), pp. 1–39.

6

[292] A. H. While, S. Marvin, and M. Kovacic. “Urban robotic experimentation: San Francisco, Tokyo and Dubai”. In: *Urban Studies* 58.4 (2021), pp. 769–786.

[293] C. Johnson and B. Kuipers. “Socially-aware navigation using topological maps and social norm learning”. In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. 2018, pp. 151–157.

[294] J. Ye, X. Li, X. Zhang, Q. Zhang, and W. Chen. “Deep learning-based human activity real-time recognition for pedestrian navigation”. In: *Sensors (Switzerland)* 20.9 (2020), pp. 1–30. ISSN: 14248220. DOI: 10.3390/s20092574.

[295] Y. Ma, H. M. Paterson, and F. E. Pollick. “A motion capture library for the study of identity, gender, and emotion perception from biological motion”. In: *Behavior research methods* 38.1 (2006), pp. 134–141.

[296] A. Rasouli, I. Kotseruba, T. Kunic, and J. K. Tsotsos. “Pie: A large-scale dataset and models for pedestrian intention estimation and trajectory prediction”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, pp. 6262–6271.

[297] V. Narayanan, B. M. Manoghar, R. P. RV, and A. Bera. “EWareNet: Emotion Aware Human Intent Prediction and Adaptive Spatial Profile Fusion for Social Robot Navigation”. In: *arXiv preprint arXiv:2011.09438* (2020).

[298] E. Senft, S. Satake, and T. Kanda. “Would You Mind Me if I Pass by You? Socially-Appropriate Behaviour for an Omni-based Social Robot in Narrow Environment”. In: 2020, pp. 539–547.

[299] Z. Chen, T. Fan, X. Zhao, J. Liang, C. Shen, H. Chen, D. Manocha, J. Pan, and W. Zhang. “Autonomous social distancing in urban environments using a quadruped robot”. In: *IEEE Access* 9 (2021), pp. 8392–8403.

[300] P. T. Singamaneni, P. Bachiller-Burgos, L. J. Manso, A. Garrell, A. Sanfeliu, A. Spalanzani, and R. Alami. “A survey on socially aware robot navigation: Taxonomy and future challenges”. In: *The International Journal of Robotics Research* (2024), p. 02783649241230562.

[301] G. Panahandeh, N. Mohammadiha, A. Leijon, and P. Händel. “Continuous hidden Markov model for pedestrian activity classification and gait analysis”. In: *IEEE Transactions on Instrumentation and Measurement* 62.5 (2013), pp. 1073–1083.

[302] D. Brščić, T. Kanda, T. Ikeda, and T. Miyashita. “Person tracking in large public spaces using 3-D range sensors”. In: *IEEE Transactions on Human-Machine Systems* 43.6 (2013), pp. 522–534.

[303] S. Oh, A. Hoogs, A. Perera, N. Cuntoor, C.-C. Chen, J. T. Lee, S. Mukherjee, J. Aggarwal, H. Lee, L. Davis, *et al.* “A large-scale benchmark dataset for event recognition in surveillance video”. In: *CVPR 2011*. IEEE. 2011, pp. 3153–3160.

[304] C. Dondrup, N. Bellotto, F. Jovan, M. Hanheide, *et al.* “Real-time multisensor people tracking for human-robot spatial interaction”. In: (2015).

[305] A. Rudenko, T. P. Kucner, C. S. Swaminathan, R. T. Chadalavada, K. O. Arras, and A. J. Lilienthal. "Thor: Human-robot navigation data collection and accurate motion trajectories dataset". In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 676–682.

[306] Z. Yan, T. Duckett, and N. Bellotto. "Online learning for human classification in 3D LiDAR-based tracking". In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 864–871.

[307] H. S. Park, J.-J. Hwang, Y. Niu, and J. Shi. "Egocentric future localization". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 4697–4705.

[308] R. Martin-Martin, M. Patel, H. Rezatofighi, A. Shenoi, J. Gwak, E. Frankel, A. Sadeghian, and S. Savarese. "Jrdb: A dataset and benchmark of egocentric robot visual perception of humans in built environments". In: *IEEE transactions on pattern analysis and machine intelligence* (2021).

[309] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool. "You'll never walk alone: Modeling social behavior for multi-target tracking". In: *2009 IEEE 12th international conference on computer vision*. IEEE. 2009, pp. 261–268.

[310] B. Majecka. "Statistical models of pedestrian behaviour in the forum". In: *Master's thesis, School of Informatics, University of Edinburgh* (2009).

[311] H. Karnan, A. Nair, X. Xiao, G. Warnell, S. Pirk, A. Toshev, J. Hart, J. Biswas, and P. Stone. "Socially compliant navigation dataset (scand): A large-scale dataset of demonstrations for social navigation". In: *IEEE Robotics and Automation Letters* 7.4 (2022), pp. 11807–11814.

[312] A. Lerner, Y. Chrysanthou, and D. Lischinski. "Crowds by example". In: *Computer graphics forum*. Vol. 26. 3. Wiley Online Library. 2007, pp. 655–664.

[313] A. Geiger, P. Lenz, and R. Urtasun. "Are we ready for autonomous driving? the kitti vision benchmark suite". In: *2012 IEEE conference on computer vision and pattern recognition*. IEEE. 2012, pp. 3354–3361.

[314] B. Benfold and I. Reid. "Stable multi-target tracking in real-time surveillance video". In: *CVPR 2011*. IEEE. 2011, pp. 3457–3464.

[315] T. Chavdarova, P. Baqué, S. Bouquet, A. Maksai, C. Jose, T. Bagautdinov, L. Lettry, P. Fua, L. Van Gool, and F. Fleuret. "Wildtrack: A multi-camera hd dataset for dense unscripted pedestrian detection". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 5030–5039.

[316] P. Kothari, S. Kreiss, and A. Alahi. "Human trajectory forecasting in crowds: A deep learning perspective". In: *IEEE Transactions on Intelligent Transportation Systems* 23.7 (2021), pp. 7386–7400.

[317] T. Randhavane, U. Bhattacharya, K. Kapsaskis, K. Gray, A. Bera, and D. Manocha. "Identifying emotions from walking using affective and deep features". In: *arXiv preprint arXiv:1906.11884* (2019).

[318] F. Loureiro, J. Avelino, P. Moreno, and A. Bernardino. “Self-perception of Interaction Errors Through Human Non-verbal Feedback and Robot Context”. In: *International Conference on Social Robotics*. Springer. 2022, pp. 475–487.

[319] M. Stiber, R. Taylor, and C.-M. Huang. “Modeling human response to robot errors for timely error detection”. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2022, pp. 676–683.

[320] E. A. Kensinger. “Remembering emotional experiences: The contribution of valence and arousal”. In: *Reviews in the Neurosciences* 15.4 (2004), pp. 241–252.

[321] F. Ahmed, A. H. Bari, and M. L. Gavrilova. “Emotion recognition from body movement”. In: *IEEE Access* 8 (2019), pp. 11761–11781.

[322] J. Vroon, Y. Zhou, and G. Kortuen. “Factors for Robot Social Navigation Among Pedestrians. [UNDER REVIEW]”. In: 2023.

[323] K. L. Koay, K. Dautenhahn, S. Woods, and M. L. Walters. “Empirical results from using a comfort level device in human-robot interaction studies”. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 2006, pp. 194–201.

[324] C. Robotics. *Husky UGV - Outdoor Field Research Robot*. 2023. URL: https://clearpathrobotics.com/husky-ugv/ (visited on 02/14/2024).

6

[325] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold Jr, M. Durbin, M. S. Edmonson, J. Fischer, D. Hymes, S. T. Kimball, *et al.* “Proxemics [and comments and replies]”. In: *Current anthropology* 9.2/3 (1968), pp. 83–108.

[326] N. Dahlbäck, A. Jönsson, and L. Ahrenberg. “Wizard of Oz studies: why and how”. In: *Proceedings of the 1st international conference on Intelligent user interfaces*. 1993, pp. 193–200.

[327] M. Airas and P. Alku. “Emotions in short vowel segments: effects of the glottal flow as reflected by the normalized amplitude quotient”. In: *Tutorial and Research Workshop on Affective Dialogue Systems*. Springer. 2004, pp. 13–24.

[328] Y. Kim and E. M. Provost. “Emotion classification via utterance-level dynamics: A pattern-based approach to characterizing affective expressions”. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. 2013, pp. 3677–3681.

[329] L. Zhang, M. M. Diraneyya, J. Ryu, C. T. Haas, and E. M. Abdel-Rahman. “Jerk as an indicator of physical exertion and fatigue”. In: *Automation in Construction* 104 (2019), pp. 120–128.

[330] C. Larboulette and S. Gibet. “A review of computable expressive descriptors of human motion”. In: *Proceedings of the 2nd International Workshop on Movement and Computing*. 2015, pp. 21–28.

[331] D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, and K. Scherer. “Toward a minimal representation of affective gestures”. In: *IEEE Transactions on Affective Computing* 2.2 (2011), pp. 106–118.

[332] Z. Hu, D. Yang, S. Cheng, L. Zhou, S. Wu, and J. Liu. “We know where they are looking at from the rgb-d camera: Gaze following in 3d”. In: *IEEE Transactions on Instrumentation and Measurement* 71 (2022), pp. 1–14.

[333] H. B. Mann and D. R. Whitney. “On a test of whether one of two random variables is stochastically larger than the other”. In: *The annals of mathematical statistics* (1947), pp. 50–60.

[334] I. M. Johnstone. “On the distribution of the largest eigenvalue in principal components analysis”. In: *The Annals of statistics* 29.2 (2001), pp. 295–327.

[335] D. L. Johanson, H. S. Ahn, and E. Broadbent. “Improving interactions with healthcare robots: a review of communication behaviours in social and healthcare contexts”. In: *International Journal of Social Robotics* 13.8 (2021), pp. 1835–1850.

[336] J. V. Gómez, N. Mavridis, and S. Garrido. “Social path planning: Generic human-robot interaction framework for robotic navigation tasks”. In: *2nd Intl. workshop on cognitive robotics systems: replicating human actions and activities*. 2013.

[337] B. Busch, J. Grizou, M. Lopes, and F. Stulp. “Learning legible motion from human–robot interactions”. In: *International Journal of Social Robotics* 9.5 (2017), pp. 765–779.

[338] R. Mirsky, X. Xiao, J. Hart, and P. Stone. “Prevention and Resolution of Conflicts in Social Navigation–a Survey”. In: *arXiv preprint arXiv:2106.12113* (2021).

[339] A. Favier, P. T. Singamaneni, and R. Alami. “An Intelligent Human Avatar to Debug and Challenge Human-aware Robot Navigation Systems”. In: *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2022, pp. 760–764.

[340] Q. Qiu, S. Yao, J. Wang, J. Ma, G. Chen, and J. Ji. “Learning to Socially Navigate in Pedestrian-rich Environments with Interaction Capacity”. In: *2022 International Conference on Robotics and Automation (ICRA)*. IEEE. 2022, pp. 279–285.

[341] P. Trautman and A. Krause. “Unfreezing the robot: Navigation in dense, interacting crowds”. In: *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2010, pp. 797–803.

[342] J. Tjomsland, S. Kalkan, and H. Gunes. “Mind your manners! a dataset and a continual learning approach for assessing social appropriateness of robot actions”. In: *Frontiers in Robotics and AI* 9 (2022), p. 669420.

[343] S. Honig and T. Oron-Gilad. “Understanding and resolving failures in human-robot interaction: Literature review and model development”. In: *Frontiers in psychology* 9 (2018), p. 861.

[344] G. Guest, A. Bunce, and L. Johnson. “How many interviews are enough? An experiment with data saturation and variability”. In: *Field methods* 18.1 (2006), pp. 59–82.

[345] R. Guldenring, M. Görner, N. Hendrich, N. J. Jacobsen, and J. Zhang. “Learning local planners for human-aware navigation in indoor environments”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2020, pp. 6053–6060.

[346] A. Ollman. “Implementation of Pose Estimation Algorithms on the Clearpath Jackal UGV”. In: (2019).

[347] R. Peddi and N. Bezzo. “An Interpretable Monitoring Framework for Virtual Physics-Based Non-Interfering Robot Social Planning”. In: *IEEE Robotics and Automation Letters* 7.2 (2022), pp. 5262–5269.

[348] J. Cohen. “A coefficient of agreement for nominal scales”. In: *Educational and psychological measurement* 20.1 (1960), pp. 37–46.

[349] S. Holm. “A simple sequentially rejective multiple test procedure”. In: *Scandinavian journal of statistics* (1979), pp. 65–70.

[350] R. Mirsky, X. Xiao, J. Hart, and P. Stone. “Conflict Avoidance in Social Navigation—a Survey”. In: *ACM Transactions on Human-Robot Interaction* 13.1 (2024), pp. 1–36.

[351] W. Wang, L. Mao, R. Wang, and B.-C. Min. “Multi-robot cooperative socially-aware navigation using multi-agent reinforcement learning”. In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2024, pp. 12353–12360.

[352] J. Nakanishi, J. Baba, I. Kuramoto, K. Ogawa, Y. Yoshikawa, and H. Ishiguro. “Smart speaker vs. social robot in a case of hotel room”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2020, pp. 11391–11396.

[353] Y. Okafuji, Y. Ozaki, J. Baba, J. Nakanishi, K. Ogawa, Y. Yoshikawa, and H. Ishiguro. “Behavioral assessment of a humanoid robot when attracting pedestrians in a mall”. In: *International journal of social robotics* 14.7 (2022), pp. 1731–1747.

[354] S. Edirisinghe, S. Satake, and T. Kanda. “Field trial of a shopworker robot with friendly guidance and appropriate admonishments”. In: *ACM Transactions on Human-Robot Interaction* 12.3 (2023), pp. 1–37.

[355] M. Van Someren, Y. F. Barnard, and J. Sandberg. “The think aloud method: a practical approach to modelling cognitive”. In: *London: AcademicPress* 11.6 (1994).

[356] K. L. Koay, M. L. Walters, and K. Dautenhahn. “Methodological issues using a comfort level device in human-robot interactions”. In: *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005*. IEEE. 2005, pp. 359–364.

[357] E. Cancelli, T. Campari, L. Serafini, A. X. Chang, and L. Ballan. “Exploiting Proximity-Aware Tasks for Embodied Social Navigation”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, pp. 10957–10967.

[358] T. Smith, Y. Chen, N. Hewitt, B. Hu, and Y. Gu. “Socially aware robot obstacle avoidance considering human intention and preferences”. In: *International journal of social robotics* 15.4 (2023), pp. 661–678.

[359] J. Feine, U. Gnewuch, S. Morana, and A. Maedche. “A taxonomy of social cues for conversational agents”. In: *International Journal of Human-Computer Studies* 132 (2019), pp. 138–161.

[360] T. Ninomiya, A. Fujita, D. Suzuki, and H. Umemuro. “Development of the Multi-dimensional Robot Attitude Scale: Constructs of people’s attitudes towards domestic robots”. In: *International Conference on Social Robotics*. Springer. 2015, pp. 482–491.

[361] A. Bera, T. Randhavane, and D. Manocha. “The Emotionally Intelligent Robot:Improving Socially-aware Human Prediction in Crowded Environments”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (2019), pp. 2–3.

[362] T. N. Canh, X. HoangVan, and N. Y. Chong. “Enhancing Social Robot Navigation with Integrated Motion Prediction and Trajectory Planning in Dynamic Human Environments”. In: *2024 24th International Conference on Control, Automation and Systems (ICCAS)*. IEEE. 2024, pp. 731–736.

[363] T.-T. Nguyen and T. D. Ngo. “Spatiotemporal motion profiles for cost-based optimal approaching pose estimation”. In: *2024 IEEE/SICE International Symposium on System Integration (SII)*. IEEE. 2024, pp. 92–98.

[364] A. I. Siam, N. F. Soliman, A. D. Algarni, F. E. Abd El-Samie, and A. Sedik. “Deploying machine learning techniques for human emotion detection”. In: *Computational intelligence and neuroscience* 2022.1 (2022), p. 8032673.

7

ACKNOWLEDGEMENTS

Completing this PhD has been a profoundly transformative journey, one that I could not have embarked upon without the support and guidance of many incredible individuals who have been with me every step of the way.

First and foremost, I express my deepest gratitude to my promoter, Professor Gerd Kortuem. From the very beginning, during our first interview, you impressed me by asking really insightful questions, highlighting that the essence of doing impactful research lies in asking the right questions. This fundamental insight has shaped my entire PhD journey. In my first year, I found it challenging to address many of your questions, which initially seemed abstract, but I gradually learned to embrace the broader vision you encouraged, focusing on the big picture rather than getting lost in minor details. Throughout my PhD, your insights have pushed me to expand the boundaries of my research, and your encouragement has been a beacon during challenging times. We have gotten along quite smoothly, and I deeply appreciate your steadfast support, especially in my last year when almost everyone else had gone. You have always been there—a reliable and trustworthy companion. My gratitude can only be a small reflection of the immense impact you've had on my academic and personal growth.

I want to extend my heartfelt thanks to Dr. Jered Vroon, who has helped me greatly during my PhD study. You are undoubtedly the smartest people I have ever seen. Initially, I held the stereotype that highly intelligent people might be difficult to approach; however, Jered, upon our first meeting, you quickly dispelled this misconception. Not only did you impress me with your intellect, but you also demonstrated immense kindness and open-mindedness, which have profoundly influenced me. Our weekly meetings have been a treasure, showcasing your intelligence and deep expertise in our field. These interactions have significantly shaped my PhD career and my future. Your kindness extended even beyond your time at TU Delft; after you left, you continued to offer me tremendous support in my PhD studies, helping me navigate through some of the toughest times. I am deeply grateful for your generosity and mentorship, and I wish you all the best in your future endeavors.

I want to express my profound gratitude to Dr. Zoltán Rusák, who was instrumental in my decision to pursue my PhD at TU Delft. You are the nicest person I have ever met and incredibly smart. Your ability to quickly grasp the essence of my ideas has always amazed me. Often, when I begin to express a thought, you immediately understand what I am trying to convey, even before I fully articulate it. Your unique quality has made our discussions extremely productive and insightful. I truly treasure all the support you have provided, which has been a cornerstone of my PhD experience. Your kindness and sharp intellect have greatly enriched my time at TU Delft, making my journey both enjoyable and academically fulfilling.

I would like to express my deepest gratitude to Chao Wang, whose support in my final year of the PhD journey has been invaluable. Chao, you have tirelessly helped me by meticulously checking my papers and improving my writing skills, never once losing patience. No matter how busy you were, you always made time to help me, and with your help, my insomnia has significantly reduced. I am truly grateful for your dedication and the impact you've had on my life. I would like to express my heartfelt gratitude to Xiaozhou Liu, who provided tremendous support during my first year. I sincerely cherish the time we've spent together. I am also greatly grateful to Qiang Liu, with whom I have had one of the smoothest and most supportive friendships throughout my PhD journey. Qiang, you have always been there for me, and I treasure our friendship deeply and appreciate all the support you've offered. Thanks also to Xinhe Yao, with whom I've shared countless memorable times. Our endless lunch meetings, where we've gossiped and discussed everything from happy moments to sad times, have been invaluable. I will never forget the afternoon we talked for hours; those moments have left a lasting impression on me. Your assistance with my experiments and career decisions has been incredibly helpful, and I am deeply thankful for that. I am grateful to Di Yan, whose friendly demeanor and tremendous support have greatly eased my PhD journey. Di, your kindness and willingness to help at every turn have not gone unnoticed, and I wish you all the best in your future endeavors. I want to thank Xueqing Miao and Tingting Wang, both of whom are exceptionally nice friends. The consistent praise I hear about you from many others only confirms your positive impact on those around you. Your kindness and support have added greatly to my PhD experience. I'd also like to thank Kangmin Mao and Heshi Guan, who have supported me greatly, and I wish you all the best.

Very personal and special thanks to our library group: TUD lib. rolling kings. This group was first initiated with only three people who usually studied in the TUD library. After more time spent in the library, more people joined. By the time I wrote the acknowledgments, there were already 11 people. We shared many wonderful memories and passed through many hard times together. I enjoy lunching with you, talking and gossiping. Special thanks to Pan Fang, one of the most important friends I have made during my PhD, and I treasure all the great times we have spent together. Thank all the members of the group: Pan Fang, Yuexin Huang, Di Yan, Xin Ren, Xiuli Wang, Zhaochong Cai, Yitao Huang, Xue Yao, Weijun Wong, and Dr. Wang.

Thanks to my parents, grandparents, uncles and aunts, and all other relatives and friends for their selfless support. I hope you all stay happy forever!

8

CURRICULUM VITÆ

Yunzhong Zhou was born in China. He graduated with a Bachelor of Mechanical Engineering and Automation from Zhejiang University of Technology in 2016. In 2019, he earned a Master of Mechanical Manufacturing and Automation from Tongji University.

Continuing his academic journey, he pursued a PhD in the Department of Sustainable Design Engineering at Delft University of Technology, starting in 2019. His research focuses on human-robot interaction, socially aware navigation, and social signal processing, with a particular interest in investigating conflicts and perceived appropriateness in mobile robots.

List of Publications

Published:

- Zhou, Y. (2023, March). Perceived Appropriateness: A Novel View for Remediating Perceived Inappropriate Robot Navigation Behaviors. In Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (pp. 781-783).
- Zhou, Y., Vroon, J., & Kortuem, G. (2024). Exploring Human Preferences for Adapting Inappropriate Robot Navigation Behaviors: A Mixed-Methods Study. IEEE Robotics and Automation Letters. [First author, see Chapter 5 of the thesis]

Under Review:

- Shared Space, Shared Interaction: A Scoping Review of Social Cues and Information Communication in Socially Aware Navigation [Second Author, see Chapter 2 of the thesis]
- Factors for Robot Social Navigation Among Pedestrians [Second Author, see Chapter 3 of the thesis]
- PARSNiP: A Novel Dataset for Better Perceived Appropriateness Detection in Robot Social Navigation using Emotion and Attention Features [First author, see Chapter 4 of the thesis]