Evaluation of Socially-Aware Robot Navigation

Yuxiang Gao * *and Chien-Ming Huang*

Department of Computer Science, The Johns Hopkins University, Baltimore, MD, United States

As mobile robots are increasingly introduced into our daily lives, it grows ever more imperative that these robots navigate with and among people in a safe and socially acceptable manner, particularly in shared spaces. While research on enabling socially-aware robot navigation has expanded over the years, there are no agreed-upon evaluation protocols or benchmarks to allow for the systematic development and evaluation of socially-aware navigation. As an effort to aid more productive development and progress comparisons, in this paper we review the evaluation methods, scenarios, datasets, and metrics commonly used in previous socially-aware navigation research, discuss the limitations of existing evaluation protocols, and highlight research opportunities for advancing socially-aware robot navigation.

1 INTRODUCTION

Fueled by advances in artificial intelligence (AI) technologies, mobile robots are realizing increased adoption in various delivery-based industries, from mail[1] and packages[2] to pizza.[3] Mobile robots designed for these consumer-facing services must not only navigate safely and efficiently to their destinations but also abide by social expectations as they move through human environments. For example, it is desirable for mobile robots to respect personal space (Althaus et al., 2004), avoid cutting through social groups (Katyal et al., 2021), move at a velocity that does not distress nearby pedestrians (Kato et al., 2015), and approach people from visible directions (Huang et al., 2014) while maintaining relevant social dynamics (Truong and Ngo, 2018). Research that investigates robot capabilities for navigating in human environments in an efficient, safe, and socially acceptable manner is commonly recognized as socially-aware navigation—also known as human-aware navigation (e.g., Kruse et al., 2013), socially compliant navigation (e.g., Kretzschmar et al., 2016), socially acceptable navigation (e.g., Shiomi et al., 2014), or socially competent navigation (e.g., Mavrogiannis et al., 2017).

While research on socially-aware navigation has expanded over the years (Kruse et al., 2013; Rios-Martinez et al., 2015; Charalampous et al., 2017; Pandey, 2017), there are no standard evaluation protocols—including methods, scenarios, datasets, and metrics—to benchmark research progress. Prior works on socially-aware robot navigation utilize a variety of evaluation protocols in custom settings, rendering comparisons of research results difficult. We argue that commonly agreed-upon evaluation protocols are key to fruitful progress, as observed in other research fields (e.g., computer

---

[1]Japan Post Co. piloted their mail delivery robot in Tokyo in October 2020.
[2]FedEx is currently developing the SameDay Bot for package delivery.
[3]Domino's launched delivery robots in Houston, TX, United States in April 2021.

FIGURE 1 | Number of publications collected by year.
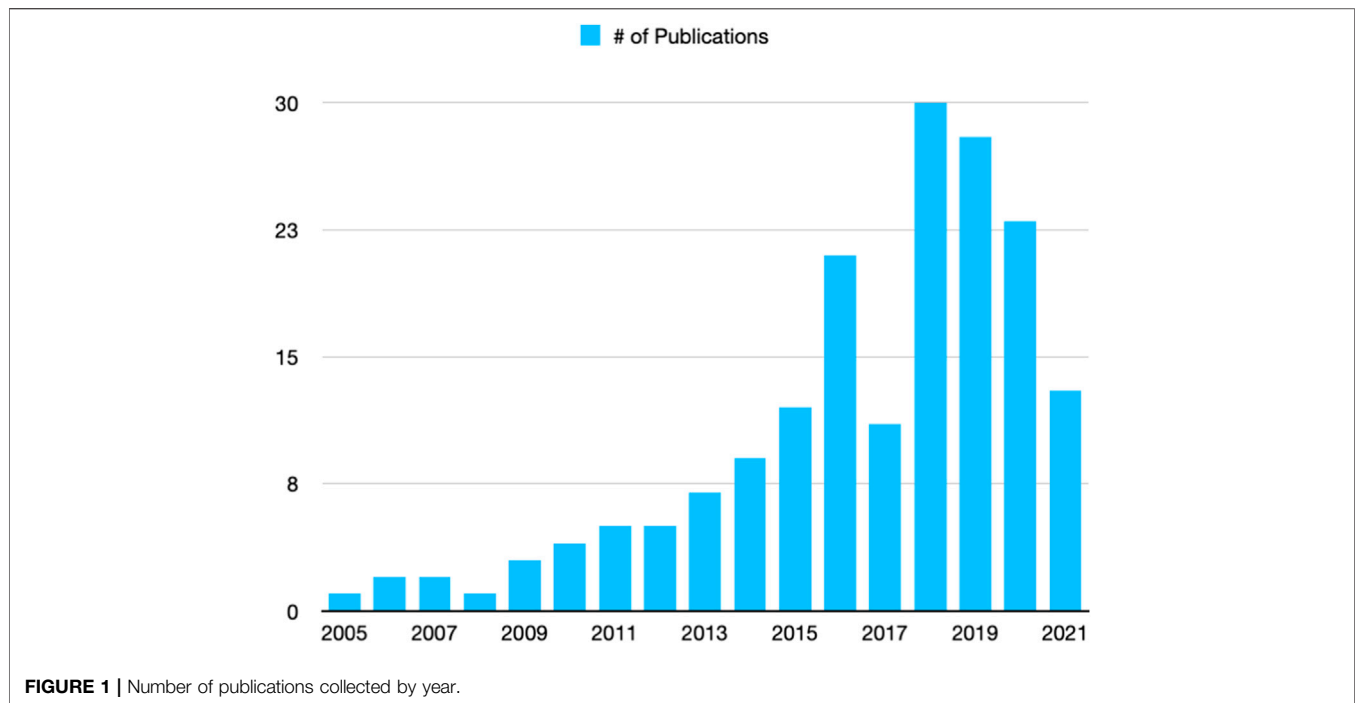
vision). As an effort to productively advance socially-aware navigation, in this paper we review commonly used evaluation methods, scenarios, datasets, and metrics in relevant prior research. We note that our review focuses on evaluation protocols rather than the algorithmic methods and systems that enable socially-aware navigation. We further note that socially-aware navigation is strongly related to an array of research topics, including human trajectory prediction, agent and crowd simulation, and robot navigation; some of the evaluation protocols reviewed in this paper may apply to these related research areas. Our review complements the recommendation for evaluation of embodied navigation suggested by Anderson et al. (2018) and can be consulted along with other general evaluation guidelines for human-robot interactions (Steinfeld et al., 2006; Young et al., 2011; Murphy and Schreckenghost, 2013).

The reminder of this paper is organized as follows. In **Section 3**, we present evaluation methods, scenarios, and datasets commonly used for evaluating socially-aware navigation. In **Section 4**, we review evaluation metrics and focus on the aspects of navigation performance, behavioral naturalness, human discomfort, and socialbility. We conclude this review with a discussion of limitations of existing evaluation protocols and opportunities for future research.

2 METHODOLOGY

Methodologically, this paper can be considered as a literature review—"*a literature review reviews published literature, implying that included materials possess some degree of permanence and, possibly, have been subject to a peer-review process. Generally, a literature review involves some process for identifying materials for potential inclusion—whether or not requiring a formal literature search—for selecting included materials, for synthesizing them in textual, tabular or graphical form and for making some analysis of their contribution or value*" (Grant and Booth, 2009). We focus on reviewing evaluation protocols for socially-aware robot navigation. While we did not follow the scoping process used for a systematic review, we identified materials (papers and datasets) for inclusion based on their relevance to the topic of socially-aware robot navigation and its evaluation methods. Specifically, we used keywords "socially-aware navigation,""socially-acceptable navigation," "human-aware navigation," or "crowd-aware navigation" when searching papers through ACM Digital Library, IEEE Xplore, and ScienceDirect. We additionally included some preprints from ArXiv through Google Scholar searches. This process yielded 188 papers in our initial search. Upon further reviewing the titles and abstracts of the papers, we removed 11 papers that did not address socially-aware robot navigation. The remaining 177 papers were published between 2005 and 2021 (**Figure 1**). A co-occurrence network of the keywords of the included papers is shown in **Figure 2**; the network illustrates three clusters that approximately represent topics related to human-robot interaction or social aspects of navigation (red), algorithmic methods for navigation (blue), and navigation systems (green). The co-occurrence network was automatically generated through Bibilometrix (Aria and Cuccurullo, 2017), a bibliometrics analysis tool, using Louvain algorithm. **Table 1** lists major venues where the 177 papers were published.

Upon collecting the 177 papers, we further reviewed the evaluation section of each paper and chose the studies that are representatives of the evaluation metrics, evaluation methods,
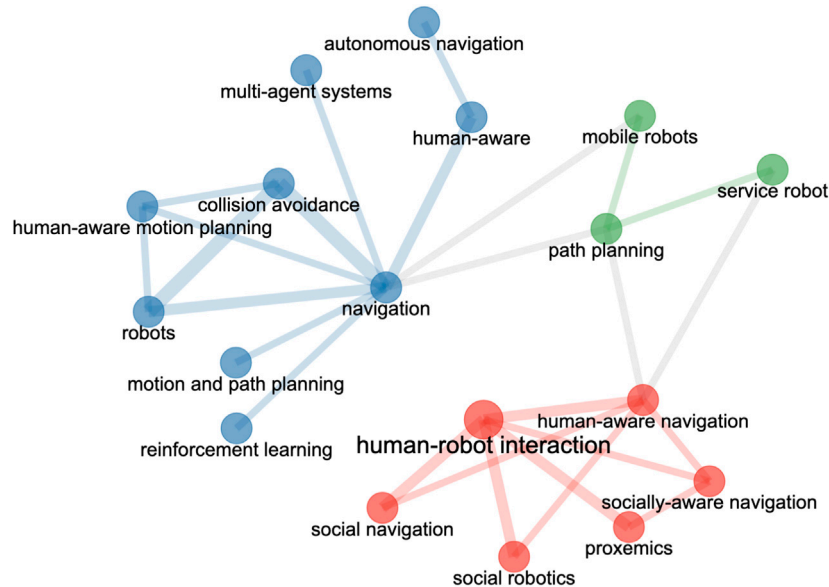
FIGURE 2 | Co-occurrence network of the keywords appeared in the collected publications. The keywords are clustered using Louvain algorithm. This graph is generated using Bibilometrix (Aria and Cuccurullo, 2017), a bibliometrics analysis tool.

TABLE 1 | Publication venues of the included 177 publications. Only venues that have more than five papers are listed.

Publication venues	#Of papers collected
IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)	21
IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)	13
IEEE International Conference on Robotics and Automation (ICRA)	12
International Journal of Social Robotics	8
IEEE Robotics and Automation Letters (RA-L)	6
ACM/IEEE International Conference on Human-Robot Interaction (HRI)	5
Others	112

datasets, and test scenarios described in the next section. Through this process, we observed that many of the evaluation metrics were originated from related works on neighboring research topics such as human trajectory prediction, autonomous robot navigation, and crowd simulation. As a result, we include relevant works on these topics to better understand the development of the evaluation methods in our report and discussion below.

3 EVALUATION METHODS, SCENARIOS, AND DATASETS

In this section, we describe evaluation methods, scenarios, and datasets commonly used in socially-aware navigation research, some of which apply directly to the problems of human trajectory prediction, crowd simulation, and general robot navigation.
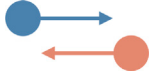
3.1 Evaluation Methods
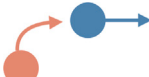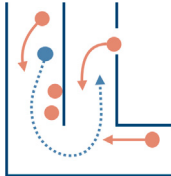
Mavrogiannis et al. (2019) classified the evaluation methods into three categories: simulation study, experimental demonstration, and experimental study. In this review, we follow a similar but more granular classification based on the type, location, and goal of the evaluation methods. Specifically, we focus on four evaluation methods—case study, simulation and demonstration, laboratory study, and field study—regularly used in socially-aware navigation research. Each method has its own advantages and disadvantages and is often used at different stages of development.

3.1.1 Case Studies

Because navigating among people in human environments involves complex, rich interactions, it is common to break down socially-aware navigation into sets of primitive, routine navigational interactions such as passing and crossing (**Table 2**). As such, prior research has utilized case studies to illustrate robot capabilities in handling these common navigational interactions. Said case studies usually involve prescribed interaction behaviors (e.g., asking the test subjects to walk in a predetermined direction or behave as if they were walking together) and environmental configurations. For example, Pacchierotti et al. (2006) studied how a person and a robot may pass each other in a hallway environment; their study involved different human behaviors,

TABLE 2 | Scenarios commonly used in evaluating socially-aware navigation. The publications that employ each scenario in simulation or real-world settings are listed respectively.

Interaction type	Illustrations	Used in simulation	Used in real-world settings
Passing		Chen et al. (2017a); Vega et al. (2019b); Yang and Peters (2019); Randhavane et al. (2019); Pandey and Alami (2010); Pérez-D'Arpino et al. (2020)	Butler and Agah (2001); Pacchierotti et al. (2006); Kretzschmar et al. (2016); Okal and Arras (2016);
Crossing		Alahi et al. (2016); Chen et al. (2017a); Chen K. et al. (2019); Sui et al. (2019); Manso et al. (2019); Khambhaita and Alami (2020); Nishimura and Yonetani (2020); Daza et al. (2021)	Guzzi et al. (2013b); Kretzschmar et al. (2016); Johnson and Kuipers (2019); Mavrogiannis et al. (2019)
Overtaking		Kirby (2010); Pandey and Alami (2010); Anvari et al. (2015); Chen et al. (2017a)	Pandey and Alami (2010); Šochman and Hogg (2011); Robicquet et al. (2016); Yang and Peters (2019)
Approaching		Turner (1981); Sisbot et al. (2005); Truong and Ngo (2018); Johnson and Kuipers (2019); Truong and Ngo (2019)	Butler and Agah (2001); Satake et al. (2009); Kato et al. (2015); Truong and Ngo (2018); Joosse et al. (2021)
Following, leading, and accompanying		Ferrer et al. (2013b); Yao et al. (2019)	Ferrer et al. (2013a); Ferrer et al. (2013b); Ferrer et al. (2017); Du et al. (2019); Repiso et al. (2020)
Combined		Okal and Arras (2014); Okal and Arras (2016); Pandey (2017); Yang and Peters (2019)	Shiomi et al. (2014); Truong and Ngo (2018); Vega et al. (2019b)

such as moving at a constant speed or stopping in the middle of the hallway, and illustrated how the robot may respond to those behaviors. Similarly, Kretzschmar et al. (2016) reported a study demonstrating how their inverse reinforcement learning approach allowed a robotic wheelchair to pass two people walking together in a hallway without cutting through the group. Truong and Ngo (2017) presented an illustrative study comparing their proactive social motion model (PSMM) against the social force model (SFM) in four experimental settings and showed that their model yielded a more socially acceptable navigation scheme. Case studies can also be presented *via* simulation; Rios-martinez et al. (2013) used a set of predefined simulated configurations of human behaviors (e.g., moving around and interacting with each other) to illustrate their proposed method for reducing discomfort caused by robot movements.

3.1.2 Simulation and Demonstrations
Simulation experiments have been regularly utilized in recent years due to advances in reinforcement learning and data-driven approaches to socially-aware navigation (e.g., Chen C. et al., 2019; Li et al., 2019; Liu Y. et al., 2020). They are particularly useful for agile development and systematic benchmarking. Simulation experiments are typically supplemented by physical demonstrations to exhibit intended robot capabilities; the objective of these demonstrations is to illustrate that the proposed algorithmic methods work not only in simulated

setups but also in the physical world with a real robot. For instance, Chen et al. (2020) first evaluated their method for crowd navigation in a simulated circle crossing scenario with five agents, after which they provided a demonstration of their method using a Pioneer robot interacting with human subjects. Katyal et al. (2020) and Liu L. et al. (2020) followed a similar method, including a simulation evaluation and a physical demonstration in their investigation of adaptive crowd navigation. Prior works that report this type of physical demonstration typically provide supplementary videos of the demonstrations (e.g., Jin et al., 2019).

Because of the popularity of simulation-based evaluation, an array of simulation platforms have been developed for robot navigation, ranging from simplistic 2D simulation [e.g., Stage (Gerkey et al., 2003) and CrowdNav (Chen C. et al., 2019), pedsimROS (Okal and Linder, 2013), MengeROS (Aroor et al., 2017)], to high-fidelity simulation leveraging existing physics and rendering engines [e.g., Webots,[4] Gibson (Xia et al., 2018), and AI2-THOR (Kolve et al., 2019)] and virtualized real environments [e.g., Matterport3D (Chang et al., 2017)]. Among these efforts, the following simulation platforms address socially-aware navigation specifically:

---

[4]https://www.cyberbotics.com

- *PedsimROS* (Okal and Linder, 2013) is a 2D simulator based on *Social Force Model* (SFM) (Helbing and Molnár, 1995). It is integrated with the ROS navigation stack and enables easy simulation of large crowds in real time.
- *MengeROS* (Aroor et al., 2017) is a 2D simulator for realistic crowd and robot simulation. It employs several backend algorithms for crowd simulation, such as *Optimal Reciprocal Collision Avoidance* (ORCA) (Van Den Berg et al., 2011), *Social Force Model* (SFM) (Helbing and Molnár, 1995), and PedVO (Curtis and Manocha, 2014).
- *CrowdNav* (Chen C. et al., 2019) is a 2D crowd and robot simulator that serves as a wrapper of OpenAI Gym (Brockman et al., 2016), which enables training and benchmarking of many reinforcement learning based algorithms.
- *SEAN-EP* (Tsoi et al., 2020) is an experimental platform for collecting human feedback on socially-aware navigation in online interactive simulations. In this web-based simulation environment, users can control a human avatar and interact with virtual robots. The platform allows for easy specification of navigation tasks and the distribution of questionnaires; it also supports simultaneous data collection from multiple participants and offloads the heavy computation of realistic simulation to cloud servers. Its web-based platform makes large-scale data collection from a diverse group of people possible.
- *SocNavBench* (Biswas et al., 2021) is another benchmark framework that aims to evaluate different socially-aware navigation methods with consistency and interpretability. As opposed to most simulation-based approaches where agent behaviors are generated from crowd simulation [e.g., using Optimal Reciprocal Collision Avoidance (ORCA) (Van Den Berg et al., 2011) or Social Force Model (SFM) (Helbing and Molnár, 1995)], human behaviors in SocNavBench are grounded in real-world datasets (i.e., UCY and ETH datasets) (**Section 3.3**). SocNavBench renders photorealistic scenes based on the trajectories recorded in these datasets and employs a set of evaluation metrics to measure path (e.g., path irregularity) and motion (e.g., average speed and energy) quality and safety (e.g., closest collision distance).
- The *CrowdBot* simulator (Grzeskowiak et al., 2021) is another benchmarking tool for socially-aware navigation that leverages the physics engine and rendering capabilities of Unity and the optimization-based Unified Microscopic Agent Navigation Simulator (UMANS) (van Toll et al., 2020) to drive the behaviors of pedestrians.

In addition to shared platforms for simulation-based evaluation, several online technical competitions have sought to benchmark socially-aware navigation. For instance, the *TrajNet++ Challenge*[5] focuses on trajectory prediction for crowded scenes and the *iGibson Challenge*[6] includes a social

---

[5] https://www.aicrowd.com/challenges/trajnet-a-trajectory-forecasting-challenge
[6] http://svl.stanford.edu/igibson/challenge.html

navigation task contextualized in indoor navigational interactions with human avatars.

3.1.3 Laboratory Studies

As opposed to case studies, which often involve prescribing human test subjects' behaviors (e.g., having them intentionally walk toward the test robot), laboratory studies utilize experimental tasks to stimulate people's natural behaviors and responses within specific contexts. Laboratory studies can be either controlled experiments or exploratory studies. Controlled experiments allow for statistical comparisons of navigation algorithms running on physical robots in semi-realistic environments; we note that controlled laboratory experiments contrast with simulation experiments, which lack the fidelity to represent real-world human-robot interactions. As an example, Mavrogiannis et al. (2019) designed an experimental task allowing three participants and a robot to move freely between six stations following a specified task procedure. A total of 105 participants were recruited for this experiment and a variety of objective and subjective metrics were collected to assess and compare three navigation strategies: Optimal Reciprocal Collision Avoidance (ORCA), Social Momentum (SM), and tele-operation. Additionally, Huang et al. (2014) evaluated how a humanoid robot may signal different levels of friendliness toward participants *via* movement behaviors—such as approach speed and direction of approach—in a mock museum setup.

Laboratory studies may also be exploratory, allowing researchers to gain early, prompt feedback from users without controlled experimentation. For instance, Bera et al. (2019) conducted an exploratory in-person lab study with 11 participants to investigate their perceptions of a robot's navigational behaviors in response to their assumed emotions.

3.1.4 Field Studies

While laboratory experiments allow for controlled comparisons, they bear reduced ecological validity; to address this limitation, field studies are used to explore people's interactions with robots in naturalistic environments. The pioneering tour guide robots RHINO (Burgard et al., 1998) and MINERVA (Thrun et al., 1999) were deployed in museums to study their collision avoidance behaviors and how people reacted to them. More recently, Satake et al. (2009) conducted a field deployment in which a mobile robot approached customers in a shopping mall to recommend shops; they explored different approach strategies and examined failed attempts. Similarly, Shiomi et al. (2014) investigated socially acceptable collision avoidance and tested their methods on a mobile robot deployed in a shopping mall for several hours with the objective of interacting with uninstructed pedestrians. Trautman et al. (2015) collected 488 runs of their experiment in a crowded cafeteria across 3 months to validate their algorithm. A benefit of deploying robots in the field is that they may reveal unexpected human behaviors; for instance, it was observed that young children "bully" a deployed mobile robot (e.g., intentionally blocking its way), which subsequently led to new research on how to recognize and avoid potential bullying behaviors in the field (Brščić et al., 2015). All in all, field studies

TABLE 3 | Datasets used in socially-aware navigation.

Name	Year	# Of people	# Of scenes	Scene type	View type	Sensor type	Annotations	Publications
UCY (Lerner et al., 2007)	**2007**	**786**	**3**	**Outdoor**	**Top-down**	**Mono**	**Trajectories, Gaze**	Lerner et al. (2007); Alahi et al. (2016); Robicquet et al. (2016); Charalampous et al. (2016); Gupta et al. (2018); Vemula et al. (2018); Amirian et al. (2019); Yao et al. (2019); Kothari et al. (2020); Liu Y. et al. (2020); Biswas et al. (2021)
ETH (Pellegrini et al., 2009)	**2009**	**750**	**2**	**Outdoor**	**Top-down**	**Mono**	**Trajectories, Group Membership**	Pellegrini et al. (2009); Alahi et al. (2016); Charalampous et al. (2016); Robicquet et al. (2016); Gupta et al. (2018); Vemula et al. (2018); Amirian et al. (2019); Yao et al. (2019); Liu Y. et al. (2020); Kothari et al. (2020); Biswas et al. (2021)
Edinburgh Informatics Forum Pedestrian Database (EIPD) (Majecka, 2009)	**2009**	**95,998**	**1**	**Outdoor**	**Top-down**	**Mono**	**Trajectories**	Majecka (2009); Luber et al. (2012); Rudenko et al. (2017)
PETS2010	**2010**	**—**	**8**	**Outdoor**	**Surveillance**	**Mono**	**—**	Bandini et al. (2014); Bastani et al. (2015); Ristani and Tomasi (2015)
VIRAT (Oh et al., 2011)	**2011**	**4,021**	**11**	**Outdoor**	**Surveillance**	**Mono**	**Trajectories**	Oh et al. (2011); Vasquez (2016)
Town Centre (Benfold and Reid, 2011)	**2011**	**230**	**1**	**Outdoor**	**Surveillance**	**Mono**	**Bounding Boxes**	Benfold and Reid (2011); Ristani and Tomasi (2015); Le and Choi (2018)
Grand Central Station (Zhou et al., 2012)	**2012**	**12,600**	**1**	**Indoor**	**Surveillance**	**Mono**	**Trajectories**	Zhou et al. (2012); Gaydashenko et al. (2018)
CFF (Alahi et al., 2014)	**2014**	**42 million**	**1**	**Outdoor**	**Top-down**	**RGB-D**	**Trajectories, Bounding Boxes**	Alahi et al. (2014); Liu et al. (2020b); Kothari et al. (2020)
Stanford Drone Dataset (Robicquet et al., 2016)	**2016**	**11,216**	**8**	**Outdoor**	**Top-down**	**Mono**	**Trajectories**	Robicquet et al. (2016); Sadeghian et al. (2018a); Sadeghian et al. (2018b); Amirian et al. (2019); Li et al. (2020)
EgoMotion (Park et al., 2016)	**2016**	**—**	**26**	**Indoor**	**FPV**	**RGB-D**	**Bounding Boxes**	Park et al. (2016)
L-CAS (Yan et al., 2017)	**2017**	**6,140**	**1**	**Indoor**	**FPV**	**RGB-D**	**Trajectories**	Yan et al. (2017); Kothari et al. (2020); Liu Y. et al. (2020)
STRAND (Hawes et al., 2017)	**2017**	**—**	**1**	**Indoor**	**FPV**	**RGB-D**	**—**	Hawes et al. (2017)
WildTrack (Chavdarova et al., 2018)	**2018**	**9,518**	**7**	**Outdoor**	**Surveillance**	**Mono**	**Trajectories, Bounding Boxes**	Chavdarova et al. (2018); Liu Y. et al. (2020); Kothari et al. (2020)
JackRabbot Dataset (Martín-Martín et al., 2021)	**2019**	**260**	**—**	**Both**	**FPV**	**RGB-D**	**Trajectories, Bounding Boxes**	Martín-Martín et al. (2021)

are difficult to execute due to the unstructured, complex nature of real-world interactions—but are vital in evaluating socially-aware navigation and may offer insights that are otherwise impossible to discover in laboratory studies.

3.2 Primitive Scenarios

In this section, we describe common primitive scenarios found in the evaluation methods discussed in the previous section. **Table 2** summarizes primitive scenarios in evaluating socially-aware navigation by the nature of the interactions involved. These scenarios include:

- Passing: This scenario captures interactions in which two agents or groups are heading in opposite directions, usually in constrained spaces such as hallways or corridors, and need to change their respective courses to pass each other.
- Crossing: This scenario captures interactions in which two agents or groups cross paths in an open space; it also considers if one of the agents or groups is stationary. Common examples of this scenario are circle crossing, where all agents are initiated on points of a circle (e.g.,

Chen C. et al., 2019; Nishimura and Yonetani, 2020), and square crossing, where all agents are initiated on the corners of a square (e.g., Guzzi et al., 2013b).
- Overtaking: This scenario captures interactions in which two agents or groups are heading in the same direction and one of them overtakes or passes the other.
- Approaching: This scenario captures interactions in which a robot intends to approach or join a stationary or moving group or individual. This scenario is observed when a robot attempts to join a static conversational group (e.g., Truong and Ngo, 2018; Yang et al., 2020), initiate an interaction (e.g., Kato et al., 2015) or follow a moving social group (e.g., Yao et al., 2019).
- Following, leading, and accompanying: This scenario captures interactions in which a robot intends to join a moving group by following (e.g., Yao et al., 2019), leading (e.g., Chuang et al., 2018), or accompanying the group side-by-side (e.g., Ferrer et al., 2017; Repiso et al., 2020).

3.3 Datasets

Table 3 details a number of datasets of human movement that are regularly used in developing algorithms for and evaluating

TABLE 4 | Evaluation metrics for navigation performance.

Metric	Description
Path Efficiency	**The ratio between the distance between two waypoints and the length of the agent's actual path between those points**
Publications: Qian et al. (2010a); Guzzi et al. (2013b); Kruse et al. (2013); Stein et al. (2016); Sebastian et al. (2017); Honig et al. (2018); Mavrogiannis et al. (2018); Neggers et al. (2018); Johnson and Kuipers (2019); Mavrogiannis et al. (2019); Vasconez et al. (2019); Ahmadi et al. (2020); Batista et al. (2020); Chadalavada et al. (2020); Liang et al. (2020); Hacinecipoglu et al. (2020); Zhang et al. (2021)	
Success Rate	**Ratio of successful trials**
Publications: Burgard et al. (1998); Jin et al. (2019); Nishimura and Yonetani (2020); Guzzi et al. (2013b); Tsai and Oh (2020); Liang et al. (2020); Chen et al. (2019b,a); Honig et al. (2018); Kamezaki et al. (2020); Qian et al. (2010b); Samsani and Muhammad (2021); Sprute et al. (2019); Yao et al. (2021)	

socially-aware navigation systems. These datasets typically capture human movement in terms of trajectories or visual bounding boxes in various indoor and outdoor environments.

The datasets are used to train models for predicting pedestrian trajectories and for generating robot movement in the presence of pedestrians. In particular, they are commonly utilized in modern data-driven approaches to socially-aware navigation, such as deep learning methods (e.g., Alahi et al., 2016; Zhou et al., 2021; Kothari et al., 2020), reinforcement learning (e.g., Chen et al., 2017a; Li et al., 2019), and generative adversarial networks (GAN) (e.g., Gupta et al., 2018; Sadeghian et al., 2018a).

Datasets are also used to evaluate and benchmark the performance of socially-aware navigation (e.g., Biswas et al., 2021; Xia et al., 2018); for example, datasets ETH (Pellegrini et al., 2009) and UCY (Lerner et al., 2007) have been widely utilized in comparing navigation baselines (e.g., Sadeghian et al., 2018a; Bisagno et al., 2019; Gupta et al., 2018). One way to use the data of human trajectories in evaluation is to replace one of the human agents with the test robot agent and compare the robot's trajectory with the corresponding prerecorded human trajectory; various evaluation metrics described in the next section may be used to quantify the differences.

4 EVALUATION METRICS

In this section, we review common metrics used to evaluate socially-aware navigation. We begin by presenting metrics for assessing navigation performance in the presence of humans. We then review metrics for representing various aspects of social compliance; in particular, we focus on the three key aspects of social compliance in socially-aware navigation as proposed by Kruse et al. (2013): naturalness—capturing motion-level similarity between robots and people; discomfort—representing the level of annoyance, stress, or danger as induced by the presence of the robot; and sociability—encapsulating how well the robot follows the social norms expected by surrounding pedestrians.

4.1 Navigation Performance

In general, prior works used navigation efficiency (Guzzi et al., 2013a; Guzzi et al., 2013b; Mavrogiannis et al., 2018; Liang et al., 2020) and *success rate* (Burgard et al., 1998; Guzzi et al., 2013b; Jin et al., 2019; Liang et al., 2020; Nishimura and Yonetani, 2020; Tsai

and Oh, 2020) to quantify the navigation performance of a robot. The common metrics for navigation performance are shown in **Table 4**.

4.1.1 Navigation Efficiency
We observed multiple measures of navigation efficiency in prior research, including path efficiency and relative throughput. Path efficiency is defined as the ratio of the distance of two waypoints to the length of the agent's actual path between those points (Mavrogiannis et al., 2019). Relative throughput (Guzzi et al., 2013b) is defined as the ratio of the number of targets the agent can reach if it ignores all collision and social constraints to the number of targets an agent can reach in an actual simulation. Both metrics calculate a ratio of performance under an ideal condition to performance under the actual condition, indicating the influences of interactions—either with people or the environment—on navigation efficiency. Other metrics for assessing efficiency include average velocity and mean time to goal (Liang et al., 2020).

4.1.2 Success Rate
In addition to the efficiency metrics discussed above, success rate is commonly used to quantify navigation performance in socially-aware navigation (Burgard et al., 1998; Guzzi et al., 2013b; Jin et al., 2019; Liang et al., 2020; Nishimura and Yonetani, 2020; Tsai and Oh, 2020). Success rate, or arrival rate, measures an agent's ability to reach its goal. When reporting success rate, it is also common to disclose the number of collisions and timeouts (e.g., Chen C. et al., 2019; Nishimura and Yonetani, 2020); a navigation trial is considered "timed out" if the agent cannot reach its goal within a specified time limit.

It is worth noting that success rate is highly dependent upon the environmental context and does not differentiate the quality of navigation between successful trials. As a result, weighted success rate metrics have been proposed to consider aspects of navigation efficiency, such as path length and completion time, while assessing success rate. These weighted metrics are single, summary metrics that represent navigation performance and can be particularly useful in reinforcement learning, which is a popular method used in recent works on robot navigation (Anderson et al., 2018; Yokoyama et al., 2021).

4.2 Behavioral Naturalness
Metrics related to naturalness focus on low-level behavioral patterns, i.e., how human-like and smooth robot movements

TABLE 5 | Evaluation metrics for naturalness.

Metric	Type		Description
	Similarity	Smoothness	
Average Displacement Error (ADE)	✓		The average L_2 distance between the predicted trajectory and the human data
Publications: Pellegrini et al. (2009); Alahi et al. (2016); Bera et al. (2017); Gupta et al. (2018); Anderson et al. (2019); Manso et al. (2019); Rudenko et al. (2019); Kothari et al. (2020); Zou et al. (2020); Hacinecipoglu et al. (2020); Zhou et al. (2021)			
Final Displacement Error (FDE)	✓		The distance between the final destination in the prediction and the human data at the same time step
Publications: Rudenko et al. (2019); Anderson et al. (2019); Kothari et al. (2020); Gupta et al. (2018); Manso et al. (2019); Bera et al. (2017); Pellegrini et al. (2009); Alahi et al. (2016); Zou et al. (2020); Zhou et al. (2021)			
Asymmetric Dynamic Time Warping	✓		A trajectory measure that doesn't require both trajectories to have the same length
Publications: Luber et al. (2012); Charalampous et al. (2016); Charalampous et al. (2017); Kostavelis et al. (2017); Avelino et al. (2021)			
Velocity and Acceleration		✓	Basic dynamics measures
Publications: Sisbot et al. (2005); Sisbot et al. (2007); Pandey and Alami (2010); Qian et al. (2010a); Qian et al. (2010b); Scandolo and Fraichard (2011); Kruse et al. (2012); Shiomi et al. (2014); Kollmitz et al. (2015); Kretzschmar et al. (2016); Truong et al. (2017); Truong and Ngo (2017); Claes and Tuyls (2018); Honig et al. (2018); Tail et al. (2018); Buchegger et al. (2019); Mavrogiannis et al. (2019); Papenmeier et al. (2019); Randhavane et al. (2019); Yoon et al. (2019); Zhong et al. (2019); Boldrer et al. (2020); Chadalavada et al. (2020); Fang et al. (2020); Hacinecipoglu et al. (2020); Ngo et al. (2020); Senft et al. (2020); Shiying et al. (2020); Gonon et al. (2021); Kivrak et al. (2021); Yao et al. (2021)			
Path Irregularity		✓	The amount of unnecessary turning over the whole path
Publications: Guzzi et al. (2013b); Mavrogiannis et al. (2018)			
Topological Complexity		✓	Measures path entanglement to quantify encounters
Publications: Kretzschmar et al. (2016); Mavrogiannis et al. (2018)			

are; measures of human similarity and path smoothness are also commonly used in human trajectory prediction research (Rudenko et al., 2019). A summary of the metrics for behavioral naturalness are shown in **Table 5**.

4.2.1 Movement Similarity
A common hypothesis in socially-aware navigation is that robots should possess navigational behaviors similar to humans' (Luber et al., 2012; Kruse et al., 2013). As a result, many prior works focus on developing and evaluating methods of producing robot trajectories that resemble those of humans under similar conditions. These prior works use a variety of measures—including displacement errors, dynamic time warping distance, and Hausdorff distance—to directly assess similarities between trajectories and end states in navigational performances.

Displacement Errors
Displacement errors are a family of metrics typically utilized in evaluating how well a predicted trajectory matches human trajectory data or a trajectory derived from other baseline methods. These metrics are widely used in pedestrian trajectory prediction research (Anderson et al., 2019; Rudenko et al., 2019; Kothari et al., 2020); they are also applied as evaluation metrics to assess the similarities between trajectories produced by navigation algorithms and by humans (Bera et al., 2017; Gupta et al., 2018; Manso et al., 2019; Kothari et al., 2020).

- Average Displacement Error (ADE) is the average L_2 distance between the predicted trajectory and the human

data to which it is being compared. It was first used to evaluate trajectory similarity in socially-aware navigation by Pellegrini et al. (2009). As the nonlinear segments of a trajectory are where most of the social interactions between a robot and pedestrians occur (Alahi et al., 2016), ADE over these nonlinear portions provides a more specific metric for assessing human-robot navigational interaction.

- Final Displacement Error (FDE) is the distance between the final destination in the predicted trajectory and the human data at the same time step. It was proposed by Alahi et al. (2016) as a complement to ADE and nonlinear ADE.

Variations such as minimum, minimum over N, best-of-N, and top n% ADE and FDE are also employed by recent pedestrian trajectory prediction works (Anderson et al., 2019; Rudenko et al., 2019); these metrics distinguish the highest accuracy a prediction can achieve on human data, which is vital for trajectory prediction. However, accuracy is not a primary concern for socially-aware navigation research, which prioritizes learning general behavior patterns rather than generating exact matches of human trajectories; therefore, these variations are rarely applied to socially-aware navigation.

Dynamic Time Warping Distance
While displacement metrics are useful in characterizing overall trajectory similarities, they are inadequate in delineating the similarities between motion behaviors at different speeds; mismatched moving speeds are especially relevant to robot navigation as mobile robots have diverse form factors, resulting in widely varying velocities when compared to humans. To address this limitation, Luber et al. (2012) took a

TABLE 6 | Evaluation metrics for human discomfort.

Metric	Type			Description	Proposed in
	Spatial	Groups	Safety		
Personal space	✓		✓	**Spatial compliance for individuals**	Hall (1966)
Publications: Pacchierotti et al. (2006); Kessler et al. (2011); Scandolo and Fraichard (2011); Torta et al. (2013); Shiomi et al. (2014); Tomari et al. (2014); Talebpour et al. (2015); Kollmitz et al. (2015); Lindner (2016); Luo and Huang (2016); Kodagoda et al. (2016); Truong and Ngo (2016); Truong et al. (2016); Truong and Ngo (2018); Forer et al. (2018); Vega-Magro et al. (2018); Fei et al. (2019); Rajamohan et al. (2019); Randhavane et al. (2019); Bachiller et al. (2021); Banisetty et al. (2021); Batista et al. (2020); Fang et al. (2020); Fuse and Tokumaru (2020); Ngo et al. (2020); Shiying et al. (2020); Vega et al. (2020); Neggers et al. (2021)					
o/p/r-space	✓	✓		**Spatial compliance for static groups**	Kendon (2010)
Publications: Kessler et al. (2011); Scandolo and Fraichard (2011); Torta et al. (2013); Shiomi et al. (2014); Tomari et al. (2014); Kollmitz et al. (2015); Talebpour et al. (2015); Batista et al. (2020); Kodagoda et al. (2016); Lindner (2016); Luo and Huang (2016); Truong et al. (2016); Truong and Ngo (2016); Fei et al. (2019); Rajamohan et al. (2019); Randhavane et al. (2019); Fang et al. (2020); Forer et al. (2018); Truong and Ngo (2018); Vega-Magro et al. (2018); Fuse and Tokumaru (2020); Ngo et al. (2020); Shiying et al. (2020); Vega et al. (2020); Bachiller et al. (2021); Banisetty et al. (2021); Neggers et al. (2021)					
Social Force Model (SFM)	✓		✓	**Measures social compliance by artificial forces**	Helbing and Molnár (1995)
Publications: Šochman and Hogg (2011); Anvari et al. (2015); Huang et al. (2018); Kivrak and Kose (2018); Yang et al. (2019); Katyal et al. (2021)					
Extended social force model	✓	✓	✓	**Adds support for social groups to SFM**	Moussaïd et al. (2010)
Publications: Yang et al. (2019); Katyal et al. (2021)					

different approach by focusing on the fact that trajectories are time-series data bearing resemblance to spoken language; they proposed a modified version of Dynamic Time Warping (Sakoe and Chiba, 1978)—an algorithm commonly used for matching spoken-word sequences at varying speeds—to transform one trajectory into another *via* time re-scaling. A dynamic time warping distance can then be calculated to compare trajectories produced by agents moving at different velocities.

4.2.2 Smoothness

The smoothness of both the geometric path and the motion profile of a robot are two important contributing factors to natural, safe navigation (Mavrogiannis et al., 2017; Mavrogiannis et al., 2018; Mavrogiannis et al., 2019). Not only are irregular paths and jittery movements inefficient, but they can also discomfort nearby pedestrians (Fraichard, 2007); therefore, it is critical to evaluate the smoothness of a robot's geometric path and motion profile in socially-aware navigation.

Path Irregularity

The smoothness of a trajectory can be characterized by the geometry of its path. For example, path irregularity (PI) (Guzzi et al., 2013b) measures the amount of unnecessary turning over the whole path a robot has traveled:

$$\text{PI} = \sum_{\text{Path}} \frac{\text{Robot Rotation} - \text{Min. Rotation Needed}}{\text{Unit Path Length}} \quad (1)$$

Topological Complexity

Prior research has also explored the use of the topological complexity index (Dynnikov and Wiest, 2007) to measure the level of entanglement in agents' paths (Mavrogiannis et al., 2018; Mavrogiannis et al., 2019). Greater path entanglement means that the agents are more likely to encounter each other during navigation, thereby inevitably forcing movement impact. Moreover, trajectories with simpler topological entanglements have been shown to be more legible (Mavrogiannis et al., 2018).

TABLE 7 | Interpersonal spaces as defined by Hall (1966).

Space name	Range	Function
Intimate space	$< .45m$	Intimate interactions
Personal space	0.45–1.2 m	Friendly interactions
Social space	1.2–3.6 m	Buffer zone for coexistence
Public space	$> 3.6m$	Public interactions

Motion Velocity and Acceleration

Velocity and acceleration are typically used to characterize motion profiles; a robot navigating in human environments is expected to keep a maximum velocity that allows it to reach the target while still maintaining a smooth acceleration profile. As an example, Mavrogiannis et al. (2019) used acceleration per segment and average energy per segment, where energy is the integral of squared velocity, to capture change in their robot's motion.

4.3 Human Discomfort

In this section, we present metrics used to measure human discomfort in socially-aware navigation. A summary of these metrics are shown in **Table 6**. We define discomfort as pedestrians' level of annoyance, stress, or danger caused by the robot's presence. Discomfort—either physical or psychological—is typically quantified by spatial models and subjective ratings (e.g., perceived safety).

4.3.1 Spatial Models
Spatial Models for Individuals

The impact of a mobile robot's navigational behavior on human comfort is difficult to quantify (Rios-martinez et al., 2013; Rios-Martinez et al., 2015; Kothari et al., 2020), as no universal "rules" are available for defining psychological comfort. Nevertheless, research suggests that the psychological comfort of humans is affected by interpersonal distance (Aiello, 1977; Baldassare, 1978; Greenberg et al., 1980). Proxemic theory (Hall, 1966) studies the

function of the space an individual maintains for different social purposes in interpersonal interactions. According to Hall's observation, an individual's perceived personal space consists of several layers of concentric circles structured by their social functions, as presented in **Table 7**; however, according to Hall, most of his subjects were healthy business professionals from the northeastern seaboard of the United States. So these spaces may vary by culture and interaction context. Other representations—such as ovoids, concentric ellipses, and asymmetric shapes—have also been used to represent personal spaces and encode more complicated social rules (Rios-Martinez et al., 2015).

Among the four spaces laid out by Hall (1966), personal space is often used as the boundary of measuring perceived safety or social comfort—either as a no-go zone, where entering the space is counted as a violation of social comfort (Rios-martinez et al., 2013; Shiomi et al., 2014), or as the boundary of a potential function that assigns costs or penalties to robots entering that space (Amaoka et al., 2009; Truong and Ngo, 2018; Yang and Peters, 2019).

However, the circular representation of personal space as suggested by Hall (1966) is quite restrictive, as it does not adequately account for characteristics of human perception and motion. As a result, many works have explored different representations to consider face orientation (Amaoka et al., 2009; Truong and Ngo, 2016), approach pose (Truong and Ngo, 2018), and motion velocity (Helbing and Molnár, 1995; Truong and Ngo, 2016). Prior research has also leveraged empirical data from experiments to model complex and realistic uses of space (Gérin-Lajoie et al., 2008; Moussaïd et al., 2009). Most notably, the Social Force Model (SFM) (Helbing and Molnár, 1995), which has been widely used to simulate human navigation behavior in social contexts, represents the constraints of personal space as attractive or repulsive forces originating from each agent. Specifically, **Eq. 2** describes how an agent i's behavior is driven by a combination of forces:

- $\overrightarrow{f}_i^{des}$: an attractive force that drives the agent to the desired goal.
- $\overrightarrow{f}_i^{obs}$: the repulsive forces from obstacles.
- $\sum_j \overrightarrow{f}_{ij}^{social}$: the sum of social repulsive forces from all other agents, j.

$$\frac{d\overrightarrow{v_i}}{dt} = \overrightarrow{f}_i^{des} + \overrightarrow{f}_i^{obs} + \sum_j \overrightarrow{f}_{ij}^{social} \qquad (2)$$

Although SFM was designed for simulating crowd behavior, it has inspired metrics seeking to quantify social comfort in socially-aware navigation. For instance, repulsive forces from obstacles and nearby agents can be used to quantify violations of social comfort and indicate "panic" behaviors in emergencies (Mehran et al., 2009). Truong and Ngo (2018) proposed the Social Individual Index (SII) to measure the physical and psychological safety of an individual. Similarly, Robicquet et al. (2016) proposed the Social Sensitivity index, which uses potential functions to model how agents interact; high social sensitivity indicates that an agent will tend to avoid other agents.

Spatial Models for Groups

The aforementioned measures consider agents individually, but we must also consider that people interact socially in group settings. Social groups can be categorized into static and dynamic groups; static groups are groups of people standing closely together and engaging in conversations as commonly seen at social events, whereas dynamic groups are groups of people walking together toward shared destinations.

Static, conversational groups can be modeled using *f-formation* (Kendon, 2010). F-formation is the spatial arrangement that group members maintain in order to respect their communal interaction space, where *o-space* is the innermost space shared by group members and reserved for in-group interactions; *p-space* surrounds the *o-space* and is the space in which members stand; and *r-space* is the outermost space separating the group from the outer world. Similar to individual discomfort, discomfort caused by a robot to a group may be measured by the robot's invasion into either the *r-space* or the *o-space*, based on the f-formation of the group (Mead et al., 2011; Rios-martinez et al., 2013; Ferrer et al., 2017).

It is commonly observed that people walk together in dynamic social groups (Federici et al., 2012; Ge et al., 2012). In addition, individual people tend to stay away from social groups when walking (Efran and Cheyne, 1973; Knowles et al., 1976; Moussaïd et al., 2010). A mobile robot deployed in human environments must know how to behave around human groups by observing such inherent etiquette. To simulate dynamic social groups, Moussaïd et al. (2010) proposed the Extended Social Force Model (ESFM).[7] As shown in **Eq. 3**, ESFM adds a new group term $\overrightarrow{f}_i^{group}$ that dictates intra-group dynamics to the original SFM. The group term, as defined by **Eq. 4**, is the summation of three forces: a cohesive force that defines attractions between group members $\overrightarrow{f}_i^{att}$; a repulsive force between group members $\overrightarrow{f}_i^{rep}$; and a gaze force $\overrightarrow{f}_i^{gaze}$ that aligns each agent with the center-of-mass of the social group, factoring in head orientation to simulate in-group social interactions.

$$\frac{d\overrightarrow{v_i}}{dt} = \overrightarrow{f}_i^{des} + \overrightarrow{f}_i^{obs} + \sum_j \overrightarrow{f}_{ij}^{social} + \overrightarrow{f}_i^{group} \qquad (3)$$

$$\overrightarrow{f}_i^{group} = \overrightarrow{f}_i^{att} + \overrightarrow{f}_i^{reb} + \overrightarrow{f}_i^{gaze} \qquad (4)$$

Similar to spatial models for individuals, spatial models for groups can be used to approximate discomfort in group interactions. As an example, to evaluate a robot's social compliance as a group member when accompanying humans, Ferrer et al. (2017) proposed a quantitative metric based on the robot's position in relation to the human members, accounting for whether or not the robot was in the field of view of the human members and the distances between group members.

4.3.2 Physical Safety

Safety is the preeminent concern in socially-aware navigation. At the most basic level, navigational safety amounts to collision avoidance: a

---

[7]Our implementation of ESFM—https://github.com/yuxiang-gao/PySocialForce

mobile robot should not have any physical contact—intentional or otherwise—with any human being. Metrics based on collision count or violation count are commonly used in simulated environments and in some robot-only experiments. For example, Liu L. et al. (2020) used the number of collisions with agents within and without the test agent's field of view, along with success rate, as the main evaluation metrics in conducting their assessment of their deep reinforcement learning based navigation algorithm in simulation. Guzzi et al. (2013b) used small-scale robots in physical experiments, allowing them to use collision count as one of their main metrics in evaluating the impact of safety margin size.

While they are arguably the most straightforward methods of measuring navigational safety violations, collision and violation counts are neither practical nor ethical to use in real-world experiments and deployments involving humans, as collisions present potential harm to the participants. Consequently, safety violations should be approximated by invasions of defined safety zones. A safety zone is typically derived from the proxemics theory proposed by Hall (1966), wherein the personal space—ranging from 0.45 to 1.2 m in Western culture—is used to measure how well a mobile robot maintains the physical safety of nearby human pedestrians (e.g., Vega et al., 2019b). Variations on safety zones are frequently used in prior works; for example, the Collision Index (CI) (Truong and Ngo, 2016), or Social Individual Index (SII) (Truong and Ngo, 2018), is a distance-based metric for capturing the violation of personal space. The index is specified in **Eq. 5**, where (x_i^p, y_i^p) is the position of the ith pedestrian p_i, (x^r, y^r) is the position of the robot, and σ_0^{px} and σ_0^{py} are the standard deviations of the personal space, empirically set to the value of 0.28:

$$\text{CI} = \max_{i=1:\, N} \exp\left(-\left(\frac{x^r - x_i^p}{\sqrt{2}\, \sigma_0^{px}} + \frac{y^r - y_i^p}{\sqrt{2}\, \sigma_0^{py}} \right) \right) \tag{5}$$

In the original definition of the index (Truong and Ngo, 2016), the standard deviations are the same for both directions ($\sigma_0^{px} = \sigma_0^{py}$), thus assuming that personal space is a perfect circle. However, as we discussed earlier, additional representations of personal space have been proposed to capture nuanced social rules, cultural influences, and specific situations (Rios-Martinez et al., 2015); therefore, this index may be adapted to account for different cultures, types of relationships, and interaction contexts by modifying the standard deviations. As another example of custom safety zones, Jin et al. (2019) defined the ego-safety zone as a circular space around an agent, analogous to the personal space, and the social-safety zone as a rectangular region stretching along an agent's current moving direction.

4.3.3 Psychological Safety
In addition to preserving physical safety, it is important to evaluate the effects of socially-aware navigation on psychological safety. Preserving psychological safety, or sometimes referred to as perceived safety, involves ensuring a stress-free and comfortable interaction (Lasota et al., 2017). Although they may not physically endanger a person, a mobile robot's navigational behaviors (e.g., how they approach and pass a person) may yet induce feelings of discomfort or stress (Butler and Agah, 2001). Consider a

situation in which a mobile robot moves rapidly toward a person and only changes its moving direction right before the imminent collision; while the robot does not make direct physical contact with the person, its navigational behavior is still likely to cause them significant stress.

A common method of assessing people's perceived psychological safety is through questionnaires. Butler and Agah (2001) asked participants to rate their comfort from 1 to 5 (with 1 being very uncomfortable and 5 being very comfortable) under different experimental conditions, including varying robot speed, distance from the human subject, and approach patterns. Similarly, Shiomi et al. (2014) used a survey to assess people's experiences interacting with a deployed mobile robot during a field study; specifically, the inquiry focused on three aspects: whether the interaction was free from obstruction, whether the person could maintain their preferred velocity in the presence of the robot, and their overall impression of the encounter.

Several established questionnaires designed for social robotics research already include questions regarding psychological safety. For example, the Godspeed questionnaire (Bartneck et al., 2008) has a sub-scale, perceived safety, comprised of questions related to subjects' relaxed/anxious, calm/agitated, and surprised/quiescent emotional states. The Robotic Social Attributes Scale (RoSAS) (Carpinella et al., 2017), based on the Godspeed questionnaire, measures people's perception and judgement of the robots' social attributes, including warmth, competence, and discomfort. The BEHAVE-II instrument (Joosse et al., 2013) includes a set of behavioral metrics that measure human responses to a robot's behavior; some of the metrics were specifically designed to gauge the discomfort caused by a robot's approach behavior (e.g., a person's step direction and step distance when a robot intrudes upon their personal space). Joosse et al. (2021) used this instrument to measure people's responses to and tolerance of personal space invasion when being approached by agents at varying speeds.

4.4 Sociability
We define sociability as a robot's conformity to complex, often nuanced, social conventions in its navigational behavior. Previously, we have described various metrics used to measure motion-level social conventions, such as approach velocity, approach pose, invasion of personal space, or passing on the dominant side (e.g., Truong and Ngo, 2016; Guzzi et al., 2013b; Yang and Peters, 2019; Pacchierotti et al., 2006). However, there exist more complex social norms around navigation-based interactions, such as elevator etiquette, waiting in a queue, asking permission to pass, and observing right-of-way at four-way intersections. A robot may move in a natural and appropriate manner that does not cause discomfort, but still violates expected, high-level social norms. For example, a robot may enter an elevator full of people in a perfectly smooth and natural fashion without first letting anyone inside leave; while the robot does not exhibit any unnaturalness or cause discomfort by violating motion-level social conventions, it breaks higher-level social norms that most people expect when riding an elevator. Measuring these high-level social norms would allow for a more holistic understanding of the impact of robot presence on humans; however, measuring sociability remains largely difficult and

is considered one of the key challenges in the field of socially-aware navigation (Mavrogiannis et al., 2021).

The Perceived Social Intelligence (PSI) scales proposed by Barchard et al. (2018); Barchard et al. (2020) evaluate 20 aspects of robotic social intelligence. For instance, the Social Competence (SOC) scale consists of four items: 1) social competence, 2) social awareness, 3) social insensitivity (reversed), and 4) strong social skills. PSI scales have been used in previous evaluations of socially-aware navigation (e.g., Barchard et al., 2020); recently, Banisetty and Williams (2021) used the perceived safety scale from the Godspeed questionnaire in conjunction with PSI to evaluate how a robot's spatial motions may communicate social norms during a pandemic via an online study. Additionally, it has been determined that robots using socially-aware navigation planners are perceived to be more socially intelligent as measured by PSI than those using traditional navigation planners (Honour et al., 2021).

In addition to using validated scales, prior research has employed custom questions relevant to specific evaluation contexts to gauge people's perceptions of robot sociability. For example, Vega et al. (2019a) used three questions—*Is the robot's behavior socially appropriate?*; *Is the robot's behavior friendly?*; and *Does the robot understand the social context and the interaction?*—to evaluate how a mobile robot may interact with people to ask for permission to pass when they block its path. All in all, how best to measure sociability remains unresolved, as opposed to the consensus on metrics for evaluating navigation performance and trajectory similarity.

5 DISCUSSION

In this paper, we review the evaluation protocols—focusing on evaluation methods, scenarios, datasets, and metrics—most commonly used in socially-aware robot navigation with the goal of facilitating further progress in this field, which currently lacks principled frameworks for development and evaluation. Prevalent evaluation methods include simulation experiments followed by experimental demonstration, as well as laboratory and field studies. Controlled experiments, either in simulation or in the physical world, typically focus on a set of primitive scenarios such as passing, crossing, and approaching. Datasets of human movements and trajectories are regularly utilized in developing and evaluating socially-aware navigation policies. Prior works have also explored a range of objective, subjective, and behavioral measures to evaluate navigation performance, naturalness of movement, physical and psychological safety, and sociability. Below, we discuss limitations of the existing evaluation protocols and open problems to solve in future research.

5.1 Limitations of Existing Evaluation Protocols
5.1.1 Evaluation Methods, Scenarios, and Datasets
Recent works on socially-aware navigation rely heavily on datasets and simulation experiments for evaluation (Mavrogiannis et al., 2021); this trend has been accelerated by

advances in reinforcement learning and data-driven approaches in general (e.g., Luber et al., 2012; Zhou et al., 2012; Alahi et al., 2014; Alahi et al., 2016; Kretzschmar et al., 2016; Park et al., 2016). However, this type of evaluation makes strong assumptions about human and robot behaviors. For example, in simulation experiments, researchers typically rely on pedestrian behavior models such as Optimal Reciprocal Collision Avoidance (ORCA) (Van Den Berg et al., 2011) (e.g., Chen et al., 2017b; Daza et al., 2021) and the Social Force Model (SFM) (Helbing and Molnár, 1995) (e.g., Katyal et al., 2021). Reciprocal behavior models such as ORCA impose the assumption that each agent is fully aware of its surroundings and the position and velocity of the other agents; this assumption of omniscience does not hold true for a real robot or person (Fraichard and Levesy, 2020). Moreover, agents trained using ORCA and SFM behave much differently than real-life agents (Mavrogiannis et al., 2021) and there exist a multitude of SFM variations (e.g., Moussaïd et al., 2009; Anvari et al., 2015; Truong and Ngo, 2017; Huang et al., 2018; Yang and Peters, 2019); therefore, it is important to ensure comparable settings for training and evaluation when comparing algorithms in simulation experiments.

To add to this concern of agent behavior assumptions, the simulators used in virtual social navigation experiments have their own limitations. While 2D simulators such as Stage (Gerkey et al., 2003) and CrowdNav (Chen C. et al., 2019) are lightweight and easy to extend, they oversimplify and abstract, rendering their results difficult to apply to the real world. Recently, several high-fidelity, photorealistic simulation environments were developed for indoor navigation, such as Matterport 3D (Chang et al., 2017) and Gibson (Xia et al., 2018). These environments offer improved simulations closer to real-world settings; however, generating realistic, grounded human social behaviors in high-fidelity simulation environments is still challenging.

Simulation experiments typically leverage datasets and metrics that quantify performance and similarity as described in **Section 4.2.1**. This reliance on datasets and quantitative metrics assumes that the human behaviors recorded in those datasets represent the optimal behaviors for a robot—despite robots possessing dynamics and dimensions largely dissimilar to humans; at best, it is highly debatable whether an exact copy of human trajectories is socially acceptable for all robots. Finally, as described in **Section 3.1**, simulation experiments are commonly followed by demonstrations with physical robots in a real-world setting; while appropriate for proofs-of-concept, these demonstrations are mainly illustrative and lack statistical rigorousness.

In contrast, laboratory studies allow for controlled experiments with statistical precision. However, such experiments are often simplistic and designed for specific navigational interactions (**Table 2**) in certain settings (e.g., passing interactions in a hallway). Moreover, it is important to note that interaction scenarios are usually evaluated out of context. Take the crossing scenario as an example; although crossing is largely evaluated in an open setting (e.g., circle crossing), people may exhibit very different crossing behaviors in real life, as shaped by their individual objectives, other

pedestrians, and the environment (e.g., in an open square or an art gallery). Furthermore, laboratory studies typically rely on convenience sampling for participant recruitment (e.g., college students and local residents), resulting in findings that may have limited generalization to a broader population.

Field studies are arguably the most challenging evaluation method to execute; they require robots to operate robustly and safely in unstructured human environments and naturally involve emergent, unprescribed human-robot interactions. While challenging and costly, field studies can provide rich, and sometimes unexpected, insights that simulation and laboratory studies cannot offer (**Section 3.1.4**).

Going forward, we predict an increased need for bridging algorithmic innovations in simulation and autonomous, real-world interactions. Deploying robots for human interaction, either in the field or in laboratory settings, will help us better understand the true limitations of robotics technology and how people experience and interact with it. We strongly advocate for more laboratory and field studies to productively advance socially-aware robot navigation and develop useful, functional mobile robots.

5.1.2 Evaluation Metrics
Navigation Performance
Socially-aware robot navigation shares many performance metrics with general robot navigation. Conventional performance metrics, such as efficiency and success rate, are commonly reported in the literature of socially-aware robot navigation. For example, path efficiency is the ratio of the optimal path's length to that of the actual path and is used to measure path disturbance to agents (either the robot or human pedestrians), while success rate measures an agent's ability to reach its goal. Though not typically used in evaluating socially-aware navigation, we believe metrics that account for both path efficiency and success rate, such as Success weighted by Path Length (SPL) (Anderson et al., 2018), Success weighted by Number of Actions (SNA) (Chen et al., 2021), and Success weighted by Completion Time (SCT) (Yokoyama et al., 2021), are useful metrics to compare navigation policies. However, these metrics should only be used for comparisons in the same setting, as different settings have different optimalities. All in all, these metrics attempt to sum up navigation trials into singular values; while such abstraction is useful for systematic comparison, it makes the assessment of fine-grained trajectory quality more difficult. To answer questions like what caused a particular defect in efficiency, researchers typically visualize trajectories for more qualitative analysis. However, it is worth noting that the most socially acceptable navigational behaviors are not necessarily efficiency- or performance-oriented.

Naturalness
A common method of measuring naturalness is quantifying the similarity between the robot's or the predicted trajectory and those observed in human data. Average Displacement Error (ADE) and Final Displacement Error (FDE) are conventional metrics for quantifying trajectory differences. Variations of displacement- or distance-based metrics may be employed to highlight certain aspects of navigation; for instance, ADE over the nonlinear portions of a trajectory may capture the effects of navigational interactions (e.g., passing and crossing). These types of metrics are typically used in benchmarking navigation algorithms against provided datasets in simulation experiments. While allowing for reproducible and systematic development and evaluation, this dataset-oriented evaluation protocol has several limitations. First, human navigational behaviors and trajectories are context-dependent. The recorded human behaviors in a dataset are specific to the scenario in which the data was collected; moreover, most datasets only include a limited number of scenarios. Therefore, the generalizability of the evaluated algorithms to different contexts is not adequately captured by these metrics. Second, robots and humans afford distinct navigational behaviors and expectations. At the physical level, robots are quite dissimilar to humans and therefore afford different navigational behaviors, such as moving speed. At the social level, it has been revealed that people exhibit different social expectations toward robots than humans; for instance, empirical data suggests that people are willing to let robots get closer to them than they let fellow humans (Joosse et al., 2021). Finally, the majority of existing datasets are limited to 2D trajectories and neglect the fact that navigational behaviors are multimodal in nature. Such limitations necessitate the inclusion of additional metrics to cover aspects of naturalness like sociability and interaction quality.

Instead of using recorded human trajectories as a gold standard for assessing naturalness, several context-independent metrics have been utilized to measure movement smoothness, which is regarded as an important indicator of naturalness. These metrics usually consider velocity and acceleration profiles and path irregularity, which captures the number of unnecessary turns in a path. However, appropriate interpretation of the results from these metrics requires reference points (e.g., is a path irregularity value of 0.72 "good?") that are difficult to obtain and may depend on various factors such as environmental context and culture.

Discomfort
Discomfort is another key dimension in which socially-aware robot navigation is evaluated; it can be characterized generally by physical and psychological safety. To approximate discomfort, prior works have relied upon spatial models including Hall (1966) theory on proxemics and personal space, *f-formation* for groups (Kendon, 2010), the Social Force Model (SFM) (Helbing and Molnár, 1995), and the Extended Social Force Model (ESFM) (Moussaïd et al., 2009). These models are particularly relevant to and useful in evaluating mobile navigation and spatial relationships; specifically, they have been adapted to define safety zones and identify abnormal behaviors (e.g., invading personal space) that may cause discomfort. For instance, prior research has used the Social Individual Index (SII), a numerical metric derived from spatial models, along with empirically determined thresholds to gauge psychological safety (Truong and Ngo, 2017). However, spatial model-based metrics are limited in several ways. First, all agents are assumed to be identical (e.g., possessing the same personal space and social

forces), neglecting individual differences observed in the real world; for instance, how people distance themselves from others depends upon personal relationships, individual characteristics, interaction contexts, and cultural norms. Second, common spatial models do not have sufficient granularity to represent environmental contexts. As an example, in SFM, repulsive forces from the environment are all treated the same; however, people move and interact differently in different contexts, and are therefore likely to have varying levels of discomfort tolerance in response to robot navigational behaviors. Third, it is difficult to encode high-level social norms (e.g., sociability) into these spatial models. Altogether, spacial model-based metrics are limited in their ability to represent, simulate, and quantify complex, nuanced social behaviors that humans expect and exhibit in navigation.

In addition to using the aforementioned metrics, discomfort may be measured by self-report ratings [e.g., the perceived safety subscale from the Godspeed questionnaire (Bartneck et al., 2008)] and behavioral indices [e.g., the BEHAVE-II instrument (Joosse et al., 2013)]. These measures are effective in revealing people's subjective experiences and genuine behavioral responses, which may not be accurately represented by objective metrics derived from spatial models. It is worth noting that these subjective and behavioral measures are collected after experiment completion and are consequently unsuitable for learning or adapting robot behavior in real time; however, some of the behavioral measures (e.g., step distance, facial expressions, and eye gaze) from BEHAVE-II may be calculated using computer vision techniques and therefore have the potential to be utilized in real-time behavioral adaptation.

Sociability

Sociability is a complex construct that characterizes a robot's conformity to high-level social conventions, which are conditioned on varying factors such as culture, interaction and environmental contexts, and individual characteristics (e.g., gender); as a result, there are no predetermined sets of high-level social conventions. Therefore, research thus far has explored social conventions that are by and large cherry-picked by the researchers themselves. For example, Pacchierotti et al. (2006) defined a set of social rules for hallway interactions, suggesting that a robot should 1) signal its intention by proactively moving to the right; 2) stay as far away from humans as the width of the hallway allows; and 3) wait until a person completely passes by before resuming normal navigation in order to avoid causing discomfort. Salek Shahrezaie et al. (2021) emphasized that social rules differ based on environmental contexts; for instance, a robot will need to behave differently in galleries, hallways, and around vending machines. The wide range of influencing factors on sociability makes it challenging to adopt a uniform evaluation standard or set of metrics. As a consequence, most prior works adopted an ad hoc approach, using custom questions to assess sociability (e.g., Vega et al., 2019a). More recently, Perceived Social Intelligence (PSI) scales (Barchard et al., 2020) offer an initial point for benchmarking the subjective construct of sociability. In order to productively advance socially-aware

navigation, however, further research is required to develop comprehensive instruments specifically designed to measure sociability and higher-level social skills in the context of navigational interactions.

5.2 Open Problems and Opportunities
5.2.1 Diverse, Dynamic Human Models and Long-Term Effects

As discussed in **Section 5.1.1**, there are several limitations to simulation-based evaluation, the most notable of which being homogeneity—all agents are driven by a static behavior engine—and omniscience—all agents have full awareness of their surroundings (Fraichard and Levesy, 2020); these assumptions are a result of the oversimplification and abstraction built into simulators. Moreover, most spatial models for crowd behavior and proxemics are derived from population data; consequently, the experiments and simulations using them often do not support a sufficiently diverse representation of different groups of people (Hurtado et al., 2021). Indeed, humans are naturally diverse and their behaviors and expectations change over time and according to complex factors like individual traits, cultures, and contexts. For example, abundant empirical evidence has demonstrated how age (e.g., Nomura et al., 2009; Flandorfer, 2012), personality (e.g., Walters et al., 2005; Robert, 2018), gender (e.g., Flandorfer, 2012; Strait et al., 2015), and cultural (e.g., Lim et al., 2020) differences may affect people's perceptions of and interactions with robots. Moreover, similar to how people gradually change their behaviors (e.g., standing closer when talking to each other) to reflect developments in a relationship (Altman and Taylor, 1973), robots must also evolve their behaviors—as opposed to exhibiting behaviors uniformly over time—to match their relationships and promote rapport with users. Not only must we develop behavior models to account for gradual changes in relationships, but we must conduct more longitudinal studies to explore how people's experiences with, perceptions of, and behaviors toward robots change over long periods of time. Buchner et al. (2013) demonstrated that a person's experience with a collaborative robot clearly changes over the course of a year; will we see similar effects in navigational human-robot interactions? Ultimately, we have three recommendations for future research:

- Enrich pedestrian models: Although there are limitations to simulation-based approaches to socially-aware navigation, these approaches allow for rapid development and systematic benchmarking and are particularly useful for early-stage validation. However, future simulation-based research must augment pedestrian models to account for human diversity; this may be achieved by including variables to represent the influencing factors we previously discussed and by introducing parameters to regulate said variables over time and according to interaction contexts.
- Examine longitudinal effects: Our understanding of the longitudinal effects of navigational human-robot interactions is fairly limited, yet such knowledge is

critical in developing and integrating mobile robots into real-life environments with the goal of interacting with and assisting people in their daily lives. As the field of socially-aware robot navigation continues to evolve, research efforts should increasingly concentrate on conducting longitudinal field studies.

- Measure and report individual characteristics: As previously mentioned, many characteristics and factors demonstrably influence general human-robot interaction. To collectively advance our understanding of navigational human-robot interaction, we encourage future works to collect and report data on individual characteristics (e.g., age, personality, gender, and culture) and how they relate to the metrics of socially-aware navigation.

5.2.2 Evaluating Mobile Robots of Different Forms

In this paper, we focus on the evaluation of socially-aware navigation in typical mobile robots that move around and interact with people in human environments, such as indoor or outdoor delivery robots. However, mobile robots can take many forms, interactions with humans can happen in different settings (e.g., where people are "on" or "inside" the robot), and human environments can include larger-scale infrastructures such as roads and highways. In particular, our review does not address two notable classes of "robot": robotic wheelchairs and autonomous vehicles. While these two categories share various characteristics in terms of socially-aware navigation, they necessitate additional evaluation considerations and methods.

Similar to traditional mobile robots, robotic wheelchairs must consider the people around them when moving through human environments (e.g., Kretzschmar et al., 2016); as such, various evaluation considerations and metrics discussed in this paper may be adapted for this category of "robot." However, robotic wheelchairs must also take into account additional considerations for their direct users; for instance, Morales et al. (2015) explored ways of including human factors (e.g., user visibility of the environment) when planning paths for a robotic wheelchair and evaluated how comfortable users felt during the ride. In support of greater accessibility and equity, more research is needed to investigate developing and evaluating methods that enable people who are robotic wheelchair-bound to engage in social interactions with individuals or groups of people (e.g., joining or following a social group) (e.g., Escobedo et al., 2014); as such, robotic wheelchairs should consider both users' and surrounding pedestrians' social signals (e.g., intent to interact). The navigation evaluation should also include behavioral indices that capture such nuanced social dynamics. Moreover, as robotic wheelchair users have varying physical disabilities, the development and evaluation of socially-aware navigation capabilities for robotic wheelchairs must pay closer attention to individual needs. Accordingly, custom metrics may be more appropriate for evaluation, as opposed to relying upon a rigid set of standardized evaluation protocols. Detailed reporting of user characteristics and specific needs would help contextualize evaluation results.

Autonomous vehicles (AVs) are up-and-coming "mobile robots" that interact with humans, including the "driver," pedestrians, and other motorists on the road. Like traditional delivery robots, AVs must drive in a safe and predictable manner, but beyond excellent safety protocols and autonomous capabilities, AVs also require critical social awareness; social interactions underlie all pedestrian-vehicle interactions (Rasouli and Tsotsos, 2020) and even AV-AV interactions are considered social coordination events (Schwarting et al., 2019). Similar to evaluating robotic wheelchair applications, the evaluation of AV technology must consider a range of stakeholders, including pedestrians (e.g., Randhavane et al., 2019; Camara et al., 2021), bicyclists (e.g., Rahman et al., 2021), and other drivers (e.g., Schwarting et al., 2019). However, AV evaluation poses additional challenges (e.g., legal regulation for high-stake, life-critical applications) and has different considerations and norms (e.g., following traffic rules). To mitigate safety concerns, recent research has leveraged modern immersive technology such as virtual reality (VR) (e.g., Goedicke et al., 2018; Mahadevan et al., 2019; Camara et al., 2021) when evaluating socially-aware AVs; for instance, Camara et al. (2021) did their user study in a virtual reality setting to evaluate pedestrians' behavior when crossing road with vehicles present. Similar to the evaluation for mobile robots, it is very important to measure the subjective perception of pedestrian-vehicle interactions (Mahadevan et al., 2019) and consider unique spatial interactions in AV applications.

To conclude, we expect to see more autonomous mobile technologies coexisting with people in their daily lives. While these technologies—ranging from mobile service robots and robotic wheelchairs to autonomous vehicles—may have domain-specific considerations for their development and evaluation, social awareness will be vital to the successful adoption of these technologies by the general population.

6 CONCLUSION

As the field of socially-aware navigation continues to evolve, it is vital to cultivate principled frameworks for the development and evaluation of mobile robots that aim to navigate in human environments in an efficient, safe, and socially acceptable manner. In this paper, we review the evaluation protocols commonly used in socially-aware robot navigation as an effort toward developing a principled evaluation framework. Our review highlights the advantages and disadvantages of different evaluation methods and metrics; in particular, while simulation experiments allow for agile development and systematic comparisons, laboratory and field studies can offer valuable insights into navigational human-robot interactions. Moreover, objective, subjective, and behavioral metrics used together offer a more comprehensive view of robot navigation performance and user experience than individual sets of metrics alone. By reviewing evaluation protocols for socially-aware robot navigation, this paper contributes to the broader vision of successful integration of socially-aware mobile technologies into our daily lives.

AUTHOR CONTRIBUTIONS

YG and C-MH co-wrote this manuscript; both authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Ahmadi, E., Meghdari, A., and Alemi, M. (2020). A Socially Aware SLAM Technique Augmented by Person Tracking Module. *J. Intell. Robot. Syst.* 99, 3–12. doi:10.1007/s10846-019-01120-z

Aiello, J. R. (1977). A Further Look at Equilibrium Theory: Visual Interaction as a Function of Interpersonal Distance. *J. Nonverbal Behav.* 1, 122–140. doi:10.1007/bf01145461

Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., and Savarese, S. (2016). "Social LSTM: Human Trajectory Prediction in Crowded Spaces," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Las Vegas: IEEE) 2016-Decem, 961–971. doi:10.1109/cvpr.2016.110

Alahi, A., Ramanathan, V., and Fei-Fei, L. (2014). Socially-aware Large-Scale Crowd Forecasting. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2211–2218. doi:10.1109/cvpr.2014.283

Althaus, P., Ishiguro, H., Kanda, T., Miyashita, T., and Christensen, H. I. (2004). "Navigation for Human-Robot Interaction Tasks," in IEEE International Conference on Robotics and Automation, 2004. Proceeding, New Orleans, LA, USA, 26 April-1 May 2004, 1894–1900. doi:10.1109/ROBOT.2004.1308100

Altman, I., and Taylor, D. A. (1973). *Social Penetration: The Development of Interpersonal Relationships*. New York: Holt Rinehart & Winston.

Amaoka, T., Laga, H., and Nakajima, M. (2009). "Modeling the Personal Space of Virtual Agents for Behavior Simulation," in 2009 International Conference on CyberWorlds 7-11 Sept. 2009,Bradford, UK 364–370. doi:10.1109/cw.2009.19

Amirian, J., Hayet, J.-B., and Pettre, J. (2019). "Social Ways: Learning Multi-Modal Distributions of Pedestrian Trajectories with GANs," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (Long Beach: IEEE), 2964–2972. doi:10.1109/CVPRW.2019.00359

Anderson, C., Du, X., Vasudevan, R., and Johnson-Roberson, M. (2019). "Stochastic Sampling Simulation for Pedestrian Trajectory Prediction," in 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Macau, China 3-8 Nov. 2019, 1–8. doi:10.1109/iros40897.2019.8967857

Anderson, P., Chang, A., Chaplot, D. S., Dosovitskiy, A., Gupta, S., Koltun, V., et al. (2018). On Evaluation of Embodied Navigation Agents. ArXiv180706757 Cs.

Anvari, B., Bell, M. G. H., Sivakumar, A., and Ochieng, W. Y. (2015). Modelling Shared Space Users via Rule-Based Social Force Model. *Transportation Res. C: Emerging Tech.* 51, 83–103. doi:10.1016/j.trc.2014.10.012

Aria, M., and Cuccurullo, C. (2017). Bibliometrix : An R-Tool for Comprehensive Science Mapping Analysis. *J. Informetrics* 11, 959–975. doi:10.1016/j.joi.2017.08.007

Aroor, A., Epstein, S. L., and Korpan, R. (2017). "MengeROS: A Crowd Simulation Tool for Autonomous Robot Navigation," in 2017 AAAI Fall Symposium Series, 3.

Avelino, J., Garcia-Marques, L., Ventura, R., and Bernardino, A. (2021). Break the Ice: A Survey on Socially Aware Engagement for Human-Robot First Encounters. *Int. J. Soc. Robotics* 13, 1851–1877. doi:10.1007/s12369-020-00720-2

Bachiller, P., Rodriguez-Criado, D., Jorvekar, R. R., Bustos, P., Faria, D. R., and Manso, L. J. (2021). A Graph Neural Network to Model Disruption in Human-Aware Robot Navigation. *Multimed. TOOLS Appl*, 1–19. doi:10.1007/s11042-021-11113-6

Baldassare, M. (1978). Human Spatial Behavior. *Annu. Rev. Sociol.* 4, 29–56. doi:10.2307/294596410.1146/annurev.so.04.080178.000333

Bandini, S., Gorrini, A., and Vizzari, G. (2014). Towards an Integrated Approach to Crowd Analysis and Crowd Synthesis: A Case Study and First Results. *Pattern Recognition Lett.* 44, 16–29. doi:10.1016/j.patrec.2013.10.003

Banisetty, S. B., Forer, S., Yliniemi, L., Nicolescu, M., and Feil-Seifer, D. (2021). Socially-aware Navigation: A Non-linear Multi-Objective Optimization Approach. ArXiv191104037 Cs. doi:10.1145/3453445

Banisetty, S. B., and Williams, T. (2021). "Implicit Communication through Social Distancing: Can Social Navigation Communicate Social Norms," in Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, 499–504.

[Dataset] Barchard, K. A., Lapping-Carr, L., Shane, R. W., Banisetty, S. B., and Feil-Seifer, D. (2018). Perceived Social Intelligence (PSI) Scales Test Manual.

Barchard, K. A., Lapping-Carr, L., Westfall, R. S., Fink-Armold, A., Banisetty, S. B., and Feil-Seifer, D. (2020). Measuring the Perceived Social Intelligence of Robots. *J. Hum.-Robot Interact.* 9, 1–29. doi:10.1145/3415139

Bartneck, C., Kulic, D., and Croft, E. (2008). Measuring the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *Tech. Rep.* 8.

Bastani, V., Campo, D., Marcenaro, L., and Regazzoni, C. (2015). "Online Pedestrian Group Walking Event Detection Using Spectral Analysis of Motion Similarity Graph," in 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 1–5. doi:10.1109/avss.2015.7301744

Batista, M. R., Macharet, D. G., and Romero, R. A. F. (2020). Socially Acceptable Navigation of People with Multi-Robot Teams. *J. Intell. Robot. Syst.* 98, 481–510. doi:10.1007/s10846-019-01080-4

Benfold, B., and Reid, I. (2011). "Stable Multi-Target Tracking in Real-Time Surveillance Video," in CVPR 2011, Colorado Springs, CO, USA, 20-25 June 2011, 3457–3464. doi:10.1109/CVPR.2011.5995667

Bera, A., Randhavane, T., and Manocha, D. (2019). "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.

Bera, A., Randhavane, T., Prinja, R., and Manocha, D. (2017). "SocioSense: Robot Navigation Amongst Pedestrians with Social and Psychological Constraints," in IEEE International Conference on Intelligent Robots and Systems, Vancouver, BC, Canada, 24-28 Sept. 2017, 7018–7025. doi:10.1109/iros.2017.8206628

Bisagno, N., Zhang, B., and Conci, N. (2019). "Group LSTM: Group Trajectory Prediction in Crowded Scenarios," in Lecture Notes in Computer Science,Computer Vision – ECCV 2018 Workshops, 213–225. doi:10.1007/978-3-030-11015-4_18

Biswas, A., Wang, A., Silvera, G., Steinfeld, A., and Admoni, H. (2021). SocNavBench: A Grounded Simulation Testing Framework for Evaluating Social Navigation. ArXiv Prepr. ArXiv210300047.

Boldrer, M., Palopoli, L., and Fontanelli, D. (2020). "Socially-aware Multi-Agent Velocity Obstacle Based Navigation for Nonholonomic Vehicles," in 2020 Ieee 44th Annual Computers, Software, and Applications Conference (Compsac 2020), 18–25. doi:10.1109/COMPSAC48688.2020.00012

Bolei Zhou, B., Xiaogang Wang, X., and Xiaoou Tang, X. (2012). "Understanding Collective Crowd Behaviors: Learning a Mixture Model of Dynamic Pedestrian-Agents," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2871–2878. doi:10.1109/cvpr.2012.6248013

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., et al. (2016). OpenAI Gym. ArXiv160601540 Cs.

Brščić, D., Kidokoro, H., Suehiro, Y., and Kanda, T. (2015). "Escaping from Children's Abuse of Social Robots," in Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (Portland: ACM), 59–66. doi:10.1145/2696454.2696468

Buchegger, K., Todoran, G., and Bader, M. (2019). "Safe and Efficient Autonomous Navigation in the Presence of Humans at Control Level," in *Advances in Service and Industrial Robotics*. (Cham: Springer), 67, 504–511. RAAD 2018. doi:10.1007/978-3-030-00232-9_53

Buchner, R., Wurhofer, D., Weiss, A., and Tscheligi, M. (2013). "Robots in Time: How User Experience in Human-Robot Interaction Changes over Time," in International Conference on Social Robotics, 138–147. doi:10.1007/978-3-319-02675-6_14

Burgard, W., Cremers, A. B., Fox, D., Hahnel, D., Lakemeyer, G., Schulz, D., et al. (1998). "The Interactive Museum Tour-Guide Robot," in Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference, AAAI 98, IAAI 98, Madison, Wisconsin, USA, July 26-30, 1998, 11–18.

Butler, J. T., and Agah, A. (2001). Psychological Effects of Behavior Patterns of a mobile Personal Robot. Auton. Robots 10, 185–202. doi:10.1023/A:1008986004181

Camara, F., Dickinson, P., and Fox, C. (2021). Evaluating Pedestrian Interaction Preferences with a Game Theoretic Autonomous Vehicle in Virtual Reality. Transportation Res. F: Traffic Psychol. Behav. 78, 410–423. doi:10.1016/j.trf.2021.02.017

Carpinella, C. M., Wyman, A. B., Perez, M. A., and Stroessner, S. J. (2017). "The Robotic Social Attributes Scale (RoSAS): Development and Validation," in 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 254–262.

Chadalavada, R. T., Andreasson, H., Schindler, M., Palm, R., and Lilienthal, A. J. (2020). Bi-directional Navigation Intent Communication Using Spatial Augmented Reality and Eye-Tracking Glasses for Improved Safety in Human-Robot Interaction. Robotics and Computer-Integrated Manufacturing 61, 101830. doi:10.1016/j.rcim.2019.101830

Chang, A., Dai, A., Funkhouser, T., Halber, M., Nießner, M., Savva, M., et al. (2017). Matterport3D: Learning from RGB-D Data in Indoor Environments. ArXiv170906158 Cs. doi:10.1109/3dv.2017.00081

Charalampous, K., Kostavelis, I., and Gasteratos, A. (2017). Recent Trends in Social Aware Robot Navigation: A Survey. Robotics Autonomous Syst. 93, 85–104. doi:10.1016/j.robot.2017.03.002

Charalampous, K., Kostavelis, I., and Gasteratos, A. (2016). Robot Navigation in Large-Scale Social Maps: An Action Recognition Approach. Expert Syst. Appl. 66, 261–273. doi:10.1016/j.eswa.2016.09.026

Chavdarova, T., Baque, P., Bouquet, S., Maksai, A., Jose, C., Bagautdinov, T., et al. (2018). WILDTRACK: A Multi-Camera HD Dataset for Dense Unscripted Pedestrian Detection. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 5030–5039. doi:10.1109/cvpr.2018.00528

Chen, C., Hu, S., Nikdel, P., Mori, G., and Savva, M. (2020). "Relational Graph Learning for Crowd Navigation," in 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Las Vegas: IEEE), 10007–10013. doi:10.1109/IROS45743.2020.9340705

Chen, C., Liu, Y., Kreiss, S., and Alahi, A. (2019). "Crowd-robot Interaction: Crowd-Aware Robot Navigation with Attention-Based Deep Reinforcement Learning," in Proceedings - IEEE International Conference on Robotics and Automation, Montreal, QC, Canada, 20-24 2019-May, 6015–6022. doi:10.1109/icra.2019.8794134

Chen, C., Majumder, S., Al-Halah, Z., Gao, R., Ramakrishnan, S. K., and Grauman, K. (2021). Learning to Set Waypoints for Audio-Visual Navigation. ArXiv200809622 Cs.

Chen, K., de Vicente, J. P., Sepulveda, G., Xia, F., Soto, A., Vazquez, M., et al. (2019). A Behavioral Approach to Visual Navigation with Graph Localization Networks. ArXiv190300445 Cs. doi:10.15607/rss.2019.xv.010

Chen, Y. F., Everett, M., Liu, M., and How, J. P. (2017a). "Socially Aware Motion Planning with Deep Reinforcement Learning," in IEEE International Conference on Intelligent Robots and Systems, Vancouver, BC, Canada, 24-28 Sept. 2017, 1343–1350. doi:10.1109/iros.2017.8202312

Chen, Y. F., Liu, M., Everett, M., and How, J. P. (2017b). "Decentralized Non-communicating Multiagent Collision Avoidance with Deep Reinforcement Learning," in Proceedings - IEEE International Conference on Robotics and Automation (Marina Bay Sands, Singapore: IEEE), 285–292. doi:10.1109/icra.2017.7989037

Chuang, T.-K., Wang, H.-C., Lin, N.-C., Chen, J.-S., Hung, C.-H., Huang, Y.-W., et al. (2018). "Deep Trail-Following Robotic Guide Dog in Pedestrian Environments for People Who Are Blind and Visually Impaired - Learning from Virtual and Real Worlds," in 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 21-25 May 2018, 5849–5855. doi:10.1109/icra.2018.8460994

Claes, D., and Tuyls, K. (2018). Multi Robot Collision Avoidance in a Shared Workspace. Auton. Robot 42, 1749–1770. doi:10.1007/s10514-018-9726-5

Curtis, S., and Manocha, D. (2014). "Pedestrian Simulation Using Geometric Reasoning in Velocity Space," in Pedestrian and Evacuation Dynamics 2012

(Cham: Springer International Publishing), 875–890. doi:10.1007/978-3-319-02447-9_73

Daza, M., Barrios-Aranibar, D., Diaz-Amado, J., Cardinale, Y., and Vilasboas, J. (2021). An Approach of Social Navigation Based on Proxemics for Crowded Environments of Humans and Robots. Micromachines 12, 193. doi:10.3390/mi12020193

Du, Y., Hetherington, N. J., Oon, C. L., Chan, W. P., Quintero, C. P., Croft, E., et al. (2019). "Group Surfing: A Pedestrian-Based Approach to Sidewalk Robot Navigation," in Proceedings - IEEE International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20-24 May 2019, 6518–6524. doi:10.1109/ICRA.2019.8793608

Dynnikov, I., and Wiest, B. (2007). On the Complexity of Braids. J. Eur. Math. Soc. 9, 801–840. doi:10.4171/jems/98

Efran, M. G., and Cheyne, J. A. (1973). Shared Space: The Co-operative Control of Spatial Areas by Two Interacting Individuals. Can. J. Behav. Science/Revue canadienne des Sci. du comportement 5, 201–210. doi:10.1037/h0082345

Escobedo, A., Spalanzani, A., and Laugier, C. (2014). "Using Social Cues to Estimate Possible Destinations when Driving a Robotic Wheelchair," in IEEE International Conference on Intelligent Robots and Systems, 3299–3304. doi:10.1109/iros.2014.6943021

Fang, F., Shi, M., Qian, K., Zhou, B., and Gan, Y. (2020). A Human-Aware Navigation Method for Social Robot Based on Multi-Layer Cost Map. Int. J. Intell. Robot. Appl. 4, 308–318. doi:10.1007/s41315-020-00125-4

Federici, M. L., Gorrini, A., Manenti, L., and Vizzari, G. (2012). Data Collection for Modeling and Simulation: Case Study at the university of milan-bicocca. Lecture Notes Comput. Science,Cellular Automata, 699–708. doi:10.1007/978-3-642-33350-7_7210.1007/978-3-642-33350-7_72

Fei, C. S., Fai, Y. C., Ming, E. S. L., Yong, L. T., Feng, D., and Hua, P. C. J. (2019). Neural-network Based Adaptive Proxemics-Costmap for Human-Aware Autonomous Robot Navigation. Int. J. Integr. Eng. 11, 101–111. doi:10.30880/ijie.2019.11.04.011

Ferrer, G., Garrell, A., and Sanfeliu, A. (2013a). Robot Companion: A Social-Force Based Approach with Human Awareness-Navigation in Crowded Environments. IEEE Int. Conf. Intell. Robots Syst., 1688–1694. doi:10.1109/iros.2013.6696576

Ferrer, G., Garrell, A., and Sanfeliu, A. (2013b). "Social-aware Robot Navigation in Urban Environments," in 2013 European Conference on Mobile Robots, Barcelona, Spain, 25-27 Sept. 2013, 331–336. doi:10.1109/ecmr.2013.6698863

Ferrer, G., Zulueta, A. G., Cotarelo, F. H., and Sanfeliu, A. (2017). Robot Social-Aware Navigation Framework to Accompany People Walking Side-By-Side. Auton. Robot 41, 775–793. doi:10.1007/s10514-016-9584-y

Flandorfer, P. (2012). Population Ageing and Socially Assistive Robots for Elderly Persons: The Importance of Sociodemographic Factors for User Acceptance. Int. J. Popul. Res. 2012, 1–13. doi:10.1155/2012/829835

Forer, S., Banisetty, S. B., Yliniemi, L., Nicolescu, M., and Feil-Seifer, D. (2018). Socially-aware Navigation Using Non-linear Multi-Objective Optimization. IEEE Int. Conf. Intell. Robots Syst. 1, 8126–8133. doi:10.1109/iros.2018.8593825

Fraichard, T. (2007). "A Short Report about Motion Safety," in IEEE Int. Conf. on Robotics and Automation, 10–14.

Fraichard, T., and Levesy, V. (2020). From Crowd Simulation to Robot Navigation in Crowds. IEEE Robot. Autom. Lett. 5, 729–735. doi:10.1109/LRA.2020.2965032

Fuse, Y., and Tokumaru, M. (2020). Navigation Model for a Robot as a Human Group Member to Adapt to Changing Conditions of Personal Space. JACIII 24, 621–629. doi:10.20965/jaciii.2020.p0621

Gaydashenko, A., Kudenko, D., and Shpilman, A. (2018). "A Comparative Evaluation of Machine Learning Methods for Robot Navigation through Human Crowds," in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 553–557. doi:10.1109/ICMLA.2018.00089

Gérin-Lajoie, M., Richards, C. L., Fung, J., and McFadyen, B. J. (2008). Characteristics of Personal Space during Obstacle Circumvention in Physical and Virtual Environments. Gait & Posture 27, 239–247. doi:10.1016/j.gaitpost.2007.03.015

Gerkey, B. P., Vaughan, R. T., and Howard, A. (2003). The Player/Stage Project: Tools for Multi-Robot and Distributed Sensor Systems. Coimbra: IEEE, 7.

Goedicke, D., Li, J., Evers, V., and Ju, W. (2018). "VR-OOM," in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal: ACM), 1–11. doi:10.1145/3173574.3173739

Gonon, D. J., Paez-Granados, D., and Billard, A. (2021). Reactive Navigation in Crowds for Non-holonomic Robots with Convex Bounding Shape. *IEEE Robot. Autom. Lett.* 6, 4728–4735. doi:10.1109/LRA.2021.3068660

Grant, M. J., and Booth, A. (2009). A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies. *Health Inf. Libr. J.* 26, 91–108. doi:10.1111/j.1471-1842.2009.00848.x

Greenberg, C. I., Strube, M. J., and Myers, R. A. (1980). A Multitrait-Multimethod Investigation of Interpersonal Distance. *J. Nonverbal Behav.* 5, 104–114. doi:10.1007/bf00986513

Grzeskowiak, F., Gonon, D., Dugas, D., Paez-Granados, D., Chung, J., Nieto, J., et al. (2021). Crowd against the Machine: A Simulation-Based Benchmark Tool to Evaluate and Compare Robot Capabilities to Navigate a Human Crowd. *ArXiv210414177 Cs.* doi:10.1109/icra48506.2021.9561694

Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., and Alahi, A. (2018). "Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2255–2264. doi:10.1109/cvpr.2018.00240

Guzzi, J., Giusti, A., Gambardella, L. M., and Di Caro, G. A. (2013a). "Local Reactive Robot Navigation: A Comparison between Reciprocal Velocity Obstacle Variants and Human-like Behavior," in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2622–2629. doi:10.1109/IROS.2013.6696726

Guzzi, J., Giusti, A., Gambardella, L. M., Theraulaz, G., and Di Caro, G. A. (2013b). "Human-friendly Robot Navigation in Dynamic Environments," in 2013 IEEE International Conference on Robotics and Automation (Karlsruhe: IEEE), 423–430. doi:10.1109/icra.2013.6630610

Hacinecipoglu, A., Konukseven, E. I., and Koku, A. B. (2020). Multiple Human Trajectory Prediction and Cooperative Navigation Modeling in Crowded Scenes. *Intel Serv. Robotics* 13, 479–493. doi:10.1007/s11370-020-00333-8

Hall, E. T. (1966). *The Hidden Dimension*. Garden City, NY: Doubleday.

Hawes, N., Burbridge, C., Jovan, F., Kunze, L., Lacerda, B., Mudrova, L., et al. (2017). The STRANDS Project: Long-Term Autonomy in Everyday Environments. *IEEE Robot. Automat. Mag.* 24, 146–156. doi:10.1109/MRA.2016.2636359

Helbing, D., and Molnár, P. (1995). Social Force Model for Pedestrian Dynamics. *Phys. Rev. E* 51, 4282–4286. doi:10.1103/physreve.51.4282

Honig, S. S., Oron-Gilad, T., Zaichyk, H., Sarne-Fleischmann, V., Olatunji, S., and Edan, Y. (2018). Toward Socially Aware Person-Following Robots. *IEEE Trans. Cogn. Dev. Syst.* 10, 936–954. doi:10.1109/tcds.2018.2825641

Honour, A., Banisetty, S. B., and Feil-Seifer, D. (2021). "Perceived Social Intelligence as Evaluation of Socially Navigation," in Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (Boulder: ACM), 519–523. doi:10.1145/3434074.3447226

Huang, C.-M., Iio, T., Satake, S., and Kanda, T. (2014). "Modeling and Controlling Friendliness for an Interactive Museum Robot," in *Robotics: Science and Systems*. Zaragoza: RSS, 12–16. doi:10.15607/rss.2014.x.025

Huang, L., Gong, J., Li, W., Xu, T., Shen, S., Liang, J., et al. (2018). Social Force Model-Based Group Behavior Simulation in Virtual Geographic Environments. *ISPRS Int. J. Geo-inf.* 7. doi:10.3390/ijgi7020079

Hurtado, J. V., Londoño, L., and Valada, A. (2021). From Learning to Relearning: A Framework for Diminishing Bias in Social Robot Navigation. *Front. Robot. AI* 8, 650325. doi:10.3389/frobt.2021.650325

Jin, J., Nguyen, N. M., Sakib, N., Graves, D., Yao, H., and Jagersand, M. (2019). Mapless Navigation Among Dynamics with Social-Safety-Awareness: A Reinforcement Learning Approach from 2D Laser Scans.

Johnson, C., and Kuipers, B. (2018). "Socially-aware Navigation Using Topological Maps and Social Norm Learning," in AIES '18 Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, 151–157. doi:10.1145/3278721.3278772

Joosse, M., Lohse, M., Berkel, N. V., Sardar, A., and Evers, V. (2021). Making Appearances. *J. Hum.-Robot Interact.* 10, 1–24. doi:10.1145/3385121

Joosse, M., Sardar, A., Lohse, M., and Evers, V. (2013). BEHAVE-II: The Revised Set of Measures to Assess Users' Attitudinal and Behavioral Responses to a Social Robot. *Int. J. Soc. Robotics* 5, 379–388. doi:10.1007/s12369-013-0191-1

Kamezaki, M., Kobayashi, A., Yokoyama, Y., Yanagawa, H., Shrestha, M., and Sugano, S. (2020). A Preliminary Study of Interactive Navigation Framework with Situation-Adaptive Multimodal Inducement: Pass-By Scenario. *Int. J. Soc. Robotics* 12, 567–588. doi:10.1007/s12369-019-00574-3

Kato, Y., Kanda, T., and Ishiguro, H. (2015). "May I Help You," in Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (Portland: Association for Computing Machinery), 35–42. HRI '15. doi:10.1145/2696454.2696463

Katyal, K. D., Popek, K., Hager, G. D., Wang, I.-J., and Huang, C.-M. (2020). "Prediction-based Uncertainty Estimation for Adaptive Crowd Navigation," in International Conference on Human-Computer Interaction, 353–368. doi:10.1007/978-3-030-50334-5_24

Katyal, K., Gao, Y., Markowitz, J., Pohland, S., Rivera, C., Wang, I.-J., et al. (2021). Learning a Group-Aware Policy for Robot Navigation. *ArXiv201212291 Cs.*

Kendon, A. (2010). "Spacing and Orientation in Co-present Interaction," in *Development of Multimodal Interfaces: Active Listening and Synchrony* (Berlin: Springer), 1–15. doi:10.1007/978-3-642-12397-9_1

Kessler, J., Schroeter, C., and Gross, H.-M. (2011). "Approaching a Person in a Socially Acceptable Manner Using a Fast Marching Planner," in *Intelligent Robotics and Applications, Pt Ii. Vol. 7102 of Lecture Notes in Artificial Intelligence*. Aachen: Springer, 368–377. doi:10.1007/978-3-642-25489-5_36

Khambhaita, H., and Alami, R. (2020). Viewing Robot Navigation in Human Environment as a Cooperative Activity. *Robotics Res.*, 285–300. doi:10.1007/978-3-030-28619-4_25

Kirby, R. (2010). *Social Robot Navigation*. Pittsburgh: ProQuest, 3470165, 232. ProQuest Diss. Theses.

Kivrak, H., Cakmak, F., Kose, H., and Yavuz, S. (2021). Social Navigation Framework for Assistive Robots in Human Inhabited Unknown Environments. *Eng. Sci. Technol. Int. J.* 24, 284–298. doi:10.1016/j.jestch.2020.08.008

Kivrak, H., and Kose, H. (2018). "Social Robot Navigation in Human-Robot Interactive Environments: Social Force Model Approach," in 2018 26th Signal Processing and Communications Applications Conference (SIU) (Cesme-Izmir, Turkey: IEEE), 1–4. doi:10.1109/siu.2018.8404328

Knowles, E. S., et al b., Haas, S., Hyde, M., and Schuchart, G. E. (1976). Group Size and the Extension of Social Space Boundaries. *J. Personal. Soc. Psychol.* 33, 647–654. doi:10.1037/0022-3514.33.5.647

Kodagoda, S., Sehestedt, S., and Dissanayake, G. (2016). Socially Aware Path Planning for mobile Robots. *ROBOTICA* 34, 513–526. doi:10.1017/S0263574714001611

Kollmitz, M., Hsiao, K., Gaa, J., and Burgard, W. (2015). "Time Dependent Planning on a Layered Social Cost Map for Human-Aware Robot Navigation," in 2015 European Conference on Mobile Robots, ECMR 2015 - Proceedings. doi:10.1109/ecmr.2015.7324184

Kolve, E., Mottaghi, R., Han, W., VanderBilt, E., Weihs, L., Herrasti, A., et al. (2019). AI2-THOR: An Interactive 3D Environment for Visual AI. *ArXiv171205474 Cs.*

Kostavelis, I., Kargakos, A., Giakoumis, D., and Tzovaras, D. (2017). "Robot's Workspace Enhancement with Dynamic Human Presence for Socially-Aware Navigation," in International Conference on Computer Vision Systems, 279–288. doi:10.1007/978-3-319-68345-4_25

Kothari, P., Kreiss, S., and Alahi, A. (2020). *Human Trajectory Forecasting in Crowds: A Deep Learning Perspective*, 1–33.

Kretzschmar, H., Spies, M., Sprunk, C., and Burgard, W. (2016). Socially Compliant mobile Robot Navigation via Inverse Reinforcement Learning. *Int. J. Robotics Res.* 35, 1289–1307. doi:10.1177/0278364915619772

Kruse, T., Basili, P., Glasauer, S., and Kirsch, A. (2012). "Legible Robot Navigation in the Proximity of Moving Humans," in IEEE Workshop on Advanced Robotics and Its Social Impacts. doi:10.1109/arso.2012.6213404

Kruse, T., Pandey, A. K., Alami, R., and Kirsch, A. (2013). Human-aware Robot Navigation: A Survey. *Robotics Autonomous Syst.* 61, 1726–1743. doi:10.1016/j.robot.2013.05.007

Lasota, P. A., Fong, T., and Shah, J. A. (2017). A Survey of Methods for Safe Human-Robot Interaction. *FNT in Robotics* 5, 261–349. doi:10.1561/2300000052

Le, A. V., and Choi, J. (2018). Robust Tracking Occluded Human in Group by Perception Sensors Network System. *J. Intell. Robot Syst.* 90, 349–361. doi:10.1007/s10846-017-0667-6

Lerner, A., Chrysanthou, Y., and Lischinski, D. (2007). Crowds by Example. *Comput. Graphics Forum* 26, 655–664. doi:10.1111/j.1467-8659.2007.01089.x

Li, J., Yang, F., Tomizuka, M., and Choi, C. (2020). *EvolveGraph: Multi-Agent Trajectory Prediction with Dynamic Relational Reasoning*. Neurips, 19.

Li, K., Xu, Y., Wang, J., and Meng, M. Q.-H. (2019). "SARL∗: Deep Reinforcement Learning Based Human-Aware Navigation for Mobile Robot in Indoor Environments," in 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), 688–694. doi:10.1109/ROBIO49542.2019.8961764

Liang, J., Patel, U., Sathyamoorthy, A. J., and Manocha, D. (2020). *Realtime Collision Avoidance for mobile Robots in Dense Crowds Using Implicit Multi-Sensor Fusion and Deep Reinforcement Learning*. Ithaca: arxiv.

Lim, V., Rooksby, M., and Cross, E. S. (2020). Social Robots on a Global Stage: Establishing a Role for Culture during Human–Robot Interaction. *Int. J. Soc. Robot.* 13, 1–27. doi:10.1007/s12369-020-00710-4

Lindner, F. (2016). "How to Count Multiple Personal-Space Intrusions in Social Robot Navigation," in *What Social Robots Can and Should Do* (Amsterdam: IOS PRESS), 323–331. vol. 290 of Frontiers in Artificial Intelligence and Applications. doi:10.3233/978-1-61499-708-5-323

Liu, L., Dugas, D., Cesari, G., Siegwart, R., and Dube, R. (2020). "Robot Navigation in Crowded Environments Using Deep Reinforcement Learning," in 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Las Vegas: IEEE), 5671–5677. doi:10.1109/IROS45743.2020.9341540

Liu, Y., Yan, Q., and Alahi, A. (2020). Social NCE: Contrastive Learning of Socially-Aware Motion Representations. ArXiv201211717 Cs.

Luber, M., Spinello, L., Silva, J., and Arras, K. O. (2012). "Socially-aware Robot Navigation: A Learning Approach," in 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (Vilamoura-Algarve, Portugal: IEEE), 902–907. doi:10.1109/iros.2012.6385716

Luo, R. C., and Huang, C. (2016). "Human-aware Motion Planning Based on Search and Sampling Approach," in 2016 IEEE WORKSHOP ON ADVANCED ROBOTICS AND ITS SOCIAL IMPACTS (ARSO) (Shanghai, China: IEEE), 226–231. IEEE Workshop on Advanced Robotics and Its Social Impacts. doi:10.1109/arso.2016.7736286

Mahadevan, K., Sanoubari, E., Somanath, S., Young, J. E., and Sharlin, E. (2019). "AV-pedestrian Interaction Design Using a Pedestrian Mixed Traffic Simulator," in DIS 2019 - Proceedings of the 2019 ACM Designing Interactive Systems Conference, 475–486. doi:10.1145/3322276.3322328

Majecka, B. (2009). Statistical Models of Pedestrian Behaviour in the Forum. Masters Thesis Sch. Inform. Univ. Edinb.

Manso, L. J., Jorvekar, R. R., Faria, D. R., Bustos, P., and Bachiller, P. (2019). *Graph Neural Networks for Human-Aware Social Navigation*. Ithaca: arxiv.

Martin-Martin, R., Patel, M., Rezatofighi, H., Shenoi, A., Gwak, J., Frankel, E., et al. (2021). JRDB: A Dataset and Benchmark of Egocentric Robot Visual Perception of Humans in Built Environments. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1. doi:10.1109/TPAMI.2021.3070543

Mavrogiannis, C., Baldini, F., Wang, A., Zhao, D., Trautman, P., Steinfeld, A., et al. (2021). Core Challenges of Social Robot Navigation: A Survey. ArXiv210305668 Cs.

Mavrogiannis, C., Hutchinson, A. M., MacDonald, J., Alves-Oliveira, P., Knepper, R. A., and Knepper, R. A. (2019). "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), Daegu, Korea (South), 11-14 March 2019 (Daegu, South Korea: IEEE), 421–430. doi:10.1109/hri.2019.8673115

Mavrogiannis, C. I., Blukis, V., and Knepper, R. A. (2017). "Socially Competent Navigation Planning by Deep Learning of Multi-Agent Path Topologies," in 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 6817–6824. doi:10.1109/iros.2017.8206601

Mavrogiannis, C. I., Thomason, W. B., and Knepper, R. A. (2018). "Social Momentum," in ACM/IEEE International Conference on Human-Robot Interaction, 361–369. doi:10.1145/3171221.3171255

Mead, R., Atrash, A., and Matarić, M. J. (2011). "Proxemic Feature Recognition for Interactive Robots: Automating Metrics from the Social Sciences," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Berlin, Heidelberg: Springer), Vol. 7072 LNAI, 52–61. doi:10.1007/978-3-642-25504-5_6

Mehran, R., Oyama, A., and Shah, M. (2009). "Abnormal Crowd Behavior Detection Using Social Force Model," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20-25 June 2009 (Miami Beach: IEEE), 935–942. doi:10.1109/cvpr.2009.5206641

Morales, Y., Watanabe, A., Ferreri, F., Even, J., Ikeda, T., Shinozawa, K., et al. (2015). "Including Human Factors for Planning Comfortable Paths," in 2015 IEEE International Conference on Robotics and Automation (ICRA), 6153–6159. doi:10.1109/icra.2015.7140063

Moussaïd, M., Helbing, D., Garnier, S., Johansson, A., Combe, M., and Theraulaz, G. (2009). Experimental Study of the Behavioural Mechanisms Underlying Self-Organization in Human Crowds. *Proc. R. Soc. B.* 276, 2755–2762. doi:10.1098/rspb.2009.0405

Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., and Theraulaz, G. (2010). The Walking Behaviour of Pedestrian Social Groups and its Impact on Crowd Dynamics. *PLoS ONE* 5, e10047–7. doi:10.1371/journal.pone.0010047

Murphy, R. R., and Schreckenghost, D. (2013). "Survey of Metrics for Human-Robot Interaction," in 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 197–198. doi:10.1109/HRI.2013.6483569

Neggers, M. M. E., Cuijpers, R. H., and Ruijten, P. A. M. (2018). "Comfortable Passing Distances for Robots," in *Social Robotics* (Cham: Springer International Publishing), 431–440. Lecture Notes in Computer Science. doi:10.1007/978-3-030-05204-1_42

Neggers, M. M. E., Cuijpers, R. H., Ruijten, P. A. M., and IJsselsteijn, W. A. (2021). Determining Shape and Size of Personal Space of a Human when Passed by a Robot. *Int. J. Soc. Robotics*, 1–12. doi:10.1007/s12369-021-00805-6

Ngo, H. Q. T., Le, V. N., Thien, V. D. N., Nguyen, T. P., and Nguyen, H. (2020). Develop the Socially Human-Aware Navigation System Using Dynamic Window Approach and Optimize Cost Function for Autonomous Medical Robot. *Adv. Mech. Eng.* 12, 168781402097943. doi:10.1177/1687814020979430

Nishimura, M., and Yonetani, R. (2020). *L2B: Learning to Balance the Safety-Efficiency Trade-Off in Interactive Crowd-Aware Robot Navigation*. Ithaca: arxiv. doi:10.1109/iros45743.2020.9341519

Nomura, T., Kanda, T., Suzuki, T., and Kato, K. (2009). Age Differences and Images of Robots. *Is* 10, 374–391. doi:10.1075/is.10.3.05nom

Oh, S., Hoogs, A., Perera, A., Cuntoor, N., Chen, C.-C., Lee, J. T., et al. (2011). "A Large-Scale Benchmark Dataset for Event Recognition in Surveillance Video," in CVPR 2011 (IEEE), 3153–3160. doi:10.1109/CVPR.2011.5995586

Okal, B., and Arras, K. O. (2016). "Learning Socially Normative Robot Navigation Behaviors with Bayesian Inverse Reinforcement Learning," in 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 16-21 May 2016, 2889–2895. doi:10.1109/icra.2016.7487452

Okal, B., and Arras, K. O. (2014). "Towards Group-Level Social Activity Recognition for mobile Robots," in Proc. 3rd IROS'2014 Workshop "Assistance Serv. Robot. Hum. Environ.

[Dataset] Okal, B., and Linder, T. (2013). *Pedestrian Simulator*. Freiburg: Social Robotics Lab, University of Freiburg.

Pacchierotti, E., Christensen, H. I., and Jensfelt, P. (2006). "Embodied Social Interaction for Service Robots in Hallway Environments," in *Field and Service Robotics* (Cham: Springer), 293–304.

Pandey, A. K., and Alami, R. (2010). "A Framework towards a Socially Aware mobile Robot Motion in Human-Centered Dynamic Environment," in IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings, 5855–5860. doi:10.1109/iros.2010.5649688

Pandey, A. (2017). Mobile Robot Navigation and Obstacle Avoidance Techniques: A Review. *Int. Robot. Autom. J.* 2, 1–12. doi:10.15406/iratj.2017.02.00023

Papenmeier, F., Uhrig, M., and Kirsch, A. (2019). Human Understanding of Robot Motion: The Role of Velocity and Orientation. *Int. J. Soc. Robotics* 11, 75–88. doi:10.1007/s12369-018-0493-4

Park, H. S., Hwang, J.-J., Niu, Y., and Shi, J. (2016). "Egocentric Future Localization," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Las Vegas: IEEE), 4697–4705. doi:10.1109/CVPR.2016.508

Pellegrini, S., Ess, A., Schindler, K., and Van Gool, L. (2009). "You'll Never Walk Alone: Modeling Social Behavior for Multi-Target Tracking," in Proceedings of the IEEE International Conference on Computer Vision, 261–268. doi:10.1109/iccv.2009.5459260

Pérez-D'Arpino, C., Liu, C., Goebel, P., Martín-Martín, R., and Savarese, S. (2020). *Robot Navigation in Constrained Pedestrian Environments Using Reinforcement Learning*. Ithaca.

Qian, K., Ma, X., Dai, X., and Fang, F. (2010a). Robotic Etiquette: Socially Acceptable Navigation of Service Robots with Human Motion Pattern Learning and Prediction. *J. Bionic Eng.* 7, 150–160. doi:10.1016/S1672-6529(09)60199-2

Qian, K., Ma, X., Dai, X., and Fang, F. (2010b). Socially Acceptable Pre-collision Safety Strategies for Human-Compliant Navigation of Service Robots. *Adv. Robotics* 24, 1813–1840. doi:10.1163/016918610X527176

Rahman, M. T., Dey, K., Das, S., and Sherfinski, M. (2021). Sharing the Road with Autonomous Vehicles: A Qualitative Analysis of the Perceptions of Pedestrians and Bicyclists. *Transportation Res. Part F: Traffic Psychol. Behav.* 78, 433–445. doi:10.1016/j.trf.2021.03.008

Rajamohan, V., Scully-Allison, C., Dascalu, S., and Feil-Seifer, D. (2019). "Factors Influencing the Human Preferred Interaction distanceFactors Influencing the Human Preferred Interaction Distance," in 2019 28th IEEE International Conference on Robot and Human Interactive Communication. RO-MAN 2019. doi:10.1109/RO-MAN46459.2019.8956404

Randhavane, T., Bera, A., Kubin, E., Wang, A., Gray, K., and Manocha, D. (2019). "Pedestrian Dominance Modeling for Socially-Aware Robot Navigation," in Proceedings - IEEE International Conference on Robotics and Automation, Montreal, QC, Canada, 20-24 May 2019, 5621–5628. doi:10.1109/icra.2019.8794465

Rasouli, A., and Tsotsos, J. K. (2020). Autonomous Vehicles that Interact with Pedestrians: A Survey of Theory and Practice. *IEEE Trans. Intell. Transport. Syst.* 21, 900–918. doi:10.1109/tits.2019.2901817

Repiso, E., Garrell, A., and Sanfeliu, A. (2020). People's Adaptive Side-By-Side Model Evolved to Accompany Groups of People by Social Robots. *IEEE Robot. Autom. Lett.* 5, 2387–2394. doi:10.1109/LRA.2020.2970676

Rios-martinez, J., Renzaglia, A., Spalanzani, A., Martinelli, A., and Laugier, C. (2013). "Navigating between People: A Stochastic Optimization Approach," in 2012 IEEE International Conference on Robotics and Automation (St Paul: IEEE) 231855, 2880–2885. doi:10.1109/icra.2012.6224934

Rios-Martinez, J., Spalanzani, A., and Laugier, C. (2015). From Proxemics Theory to Socially-Aware Navigation: A Survey. *Int. J. Soc. Robotics* 7, 137–153. doi:10.1007/s12369-014-0251-1

Ristani, E., and Tomasi, C. (2015). Tracking Multiple People Online and in Real Time. *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.*, 444–459. doi:10.1007/978-3-319-16814-2_29

Robert, L. (2018). "Personality in the Human Robot Interaction Literature: A Review and Brief Critique," in *Personality in the Human Robot Interaction Literature: A Review and Brief Critique, Proceedings of the 24th Americas Conference on Information Systems*. Editor L. P. Robert. New Orleans: AMCIS, 16–18.

Robicquet, A., Sadeghian, A., Alahi, A., and Savarese, S. (2016). Learning Social Etiquette: Human Trajectory Understanding in Crowded Scenes. *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma* 9912 LNCS, 549–565. doi:10.1007/978-3-319-46484-8_33

Rudenko, A., Palmieri, L., and Arras, K. O. (2017). "Predictive Planning for a mobile Robot in Human Environments," in Proceedings of the Workshop on AI Planning and Robotics: Challenges and Methods (at ICRA 2017).

Rudenko, A., Palmieri, L., Herman, M., Kitani, K. M., Gavrila, D. M., and Arras, K. O. (2019). *Human Motion Trajectory Prediction: A Survey*. Ithaca: arxiv, 1–33.

Sadeghian, A., Kosaraju, V., Sadeghian, A., Hirose, N., Rezatofighi, S. H., and Savarese, S. (2018a). *SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical Constraints*. Long Beach: IEEE.

Sadeghian, A., Legros, F., Voisin, M., Vesel, R., Alahi, A., and Savarese, S. (2018b). CAR-net: Clairvoyant Attentive Recurrent Network. *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma* 11215 LNCS, 162–180. doi:10.1007/978-3-030-01252-6_10

Sakoe, H., and Chiba, S. (1978). Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Trans. Acoust. Speech, Signal. Process.* 26, 43–49. doi:10.1109/tassp.1978.1163055

Salek Shahrezaie, R., Banisetty, S. B., Mohammadi, M., and Feil-Seifer, D. (2021). "Towards Deep Reasoning on Social Rules for Socially Aware Navigation," in Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (Boulder: ACM), 515–518. doi:10.1145/3434074.3447225

Samsani, S. S., and Muhammad, M. S. (2021). Socially Compliant Robot Navigation in Crowded Environment by Human Behavior Resemblance Using Deep Reinforcement Learning. *IEEE Robot. Autom. Lett.* 6, 5223–5230. doi:10.1109/LRA.2021.3071954

Satake, S., Kanda, T., Glas, D. F., Imai, M., Ishiguro, H., and Hagita, N. (2009). "How to Approach Humans," in Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction - HRI '09 (La Jolla: ACM Press), 109. doi:10.1145/1514095.1514117

Scandolo, L., and Fraichard, T. (2011). "An Anthropomorphic Navigation Scheme for Dynamic Scenarios," in Proceedings - IEEE International Conference on Robotics and Automation, 809–814. doi:10.1109/icra.2011.5979772

Schwarting, W., Pierson, A., Alonso-Mora, J., Karaman, S., and Rus, D. (2019). Social Behavior for Autonomous Vehicles. *Proc. Natl. Acad. Sci. USA* 116, 24972–24978. doi:10.1073/pnas.1820676116

Sebastian, M., Banisetty, S. B., and Feil-Seifer, D. (2017). Socially-aware Navigation Planner Using Models of Human-Human Interaction. doi:10.1109/ROMAN.2017.8172334

Senft, E., Satake, S., and Kanda, T. (2020). "Would You Mind Me if I Pass by You," in Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (Cambridge, United Kingdom: ACM IEEE International Conference on Human-Robot Interaction), 539–547. doi:10.1145/3319502.3374812

Shiomi, M., Zanlungo, F., Hayashi, K., and Kanda, T. (2014). Towards a Socially Acceptable Collision Avoidance for a mobile Robot Navigating Among Pedestrians Using a Pedestrian Model. *Int. J. Soc. Robotics* 6, 443–455. doi:10.1007/s12369-014-0238-y

Sisbot, E., Alami, R., Simeon, T., Dautenhahn, K., Walters, M., Woods, S., et al. (2005). "Navigation in the Presence of Humans," in 5th IEEE-RAS International Conference on Humanoid Robots, 2005 (IEEE-RAS International Conference on Humanoid Robots), 181–188.

Sisbot, E. A., Marin-Urias, L. F., Alami, R., and Simeon, T. (2007). A Human Aware mobile Robot Motion Planner. *IEEE Trans. Robot.* 23, 874–883. doi:10.1109/TRO.2007.904911

Sochman, J., and Hogg, D. C. (2011). Who Knows Who - Inverting the Social Force Model for Finding Groups. *Proc. IEEE Int. Conf. Comput. Vis.*, 830–837. doi:10.1109/iccvw.2011.6130338

Sprute, D., Tönnies, K., and König, M. (2019). A Study on Different User Interfaces for Teaching Virtual Borders to mobile Robots. *Int. J. Soc. Robotics* 11, 373–388. doi:10.1007/s12369-018-0506-3

Stein, P., Spalanzani, A., Santos, V., and Laugier, C. (2016). Leader Following: A Study on Classification and Selection. *Robotics Autonomous Syst.* 75, 79–95. doi:10.1016/j.robot.2014.09.028

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., et al. (2006). "Common Metrics for Human-Robot Interaction," in HRI 2006: Proceedings of the 2006 ACM Conference on Human-Robot Interaction, 33–40. doi:10.1145/1121241.1121249

Strait, M., Briggs, P., and Scheutz, M. (2015). "Gender, More So Than Age, Modulates Positive Perceptions of Language-Based Human-Robot Interactions," in 4th International Symposium on New Frontiers in Human Robot Interaction, 21–22.

Sui, Z., Pu, Z., Yi, J., and Xiong, T. (2019). "Formation Control with Collision Avoidance through Deep Reinforcement Learning," in 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14-19 July 2019, 1–8. doi:10.1109/ijcnn.2019.8851906

Sun, S., Zhao, X., Li, Q., and Tan, M. (2020). Inverse Reinforcement Learning-Based Time-dependent A* Planner for Human-Aware Robot Navigation with Local Vision. *Adv. Robotics* 34, 888–901. doi:10.1080/01691864.2020.1753569

Tai, L., Zhang, J., Liu, M., and Burgard, W. (2018). "Socially Compliant Navigation through Raw Depth Inputs with Generative Adversarial Imitation Learning," in Proceedings - IEEE International Conference on Robotics and Automation, 1111–1117. doi:10.1109/icra.2018.8460968

Talebpour, Z., Navarro, I., and Martinoli, A. (2015). "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) (Nagoya, Japan: IEEE), 63–68. doi:10.1109/sii.2015.7404955

Thrun, S., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., et al. (1999). *MINERVA: A Second-Generation Museum Tour-Guide Robot*. Detroit: IEEE.

Tomari, R., Kobayashi, Y., and Kuno, Y. (2014). "Analysis of Socially Acceptable Smart Wheelchair Navigation Based on Head Cue Information, Procedia Computer Science," in *Medical and Rehabilitation Robotics and Instrumentation (Mrri2013)*, 198–205. vol. 42 of Procedia Computer Science (Amsterdam: Elsevier). doi:10.1016/j.procs.2014.11.052

Torta, E., Cuijpers, R. H., and Juola, J. F. (2013). Design of a Parametric Model of Personal Space for Robotic Social Navigation. *Int. J. Soc. Robotics* 5, 357–365. doi:10.1007/s12369-013-0188-9

Trautman, P., Ma, J., Murray, R. M., and Krause, A. (2015). Robot Navigation in Dense Human Crowds: Statistical Models and Experimental Studies of Human-Robot Cooperation. *Int. J. Robotics Res.* 34, 335–356. doi:10.1177/0278364914557874

Truong, X.-T., and Ngo, T.-D. (2018). "To Approach Humans?": A Unified Framework for Approaching Pose Prediction and Socially Aware Robot

Navigation. *IEEE Trans. Cogn. Dev. Syst.* 10, 557–572. doi:10.1109/tcds.2017.2751963

Truong, X.-T., and Ngo, T.-D. (2016). Dynamic Social Zone Based mobile Robot Navigation for Human Comfortable Safety in Social Environments. *Int. J. Soc. Robotics* 8, 663–684. doi:10.1007/s12369-016-0352-0

Truong, X.-T., and Ngo, T. D. (2017). Toward Socially Aware Robot Navigation in Dynamic and Crowded Environments: A Proactive Social Motion Model. *IEEE Trans. Automat. Sci. Eng.* 14, 1743–1760. doi:10.1109/tase.2017.2731371

Truong, X.-T., Yoong, V. N., and Ngo, T.-D. (2017). Socially Aware Robot Navigation System in Human Interactive Environments. *Intel Serv. Robotics* 10, 287–295. doi:10.1007/s11370-017-0232-y

Truong, X. T., and Ngo, T. D. (2019). "An Integrative Approach of Social Dynamic Long Short-Term Memory and Deep Reinforcement Learning for Socially Aware Robot Navigation," in Long-Term Human Motion Prediction Workshop ICRA 2019, Montreal, QC, May 20-24, 2019.

Truong, X. T., Ou, Y. S., and Ngo, T.-D. (2016). "Towards Culturally Aware Robot Navigation," in 2016 Ieee International Conference on Real-Time Computing and Robotics (Ieee Rcar) (Angkor Wat, Cambodia: IEEE), 63–69. doi:10.1109/rcar.2016.7784002

Tsai, C.-E., and Oh, J. (2020). "A Generative Approach for Socially Compliant Navigation," in 2020 IEEE International Conference on Robotics and Automation (ICRA), 2160–2166. doi:10.1109/ICRA40945.2020.9197497

Tsoi, N., Hussein, M., Fugikawa, O., Zhao, J. D., and Vázquez, M. (2020). SEAN-EP: A Platform for Collecting Human Feedback for Social Robot Navigation at Scale. ArXiv201212336 Cs.

Turner, J. C. (1981). Towards a Cognitive Redefinition of the Social Group. *Cah. Psychol. Cogn. Psychol. Cogn.* 1, 93–118.

Van Den Berg, J., Guy, S. J., Lin, M., and Manocha, D. (2011). "Reciprocal N-Body Collision Avoidance," in *Robotics Research*. Berlin: Springer, 3–19. doi:10.1007/978-3-642-19457-3_1

van Toll, W., Grzeskowiak, F., Gandía, A. L., Amirian, J., Berton, F., Bruneau, J., et al. (2020). "Generalized Microscropic Crowd Simulation Using Costs in Velocity Space," in Symposium on Interactive 3D Graphics and Games (ACM), 1–9. doi:10.1145/3384382.3384532

Vasconez, J. P., Guevara, L., and Cheein, F. A. (2019). "Social Robot Navigation Based on HRI Non-verbal Communication," in Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, 957–960. doi:10.1145/3297280.3297569

Vasquez, D. (2016). "Novel Planning-Based Algorithms for Human Motion Prediction," in 2016 IEEE International Conference on Robotics and Automation (ICRA) (Stockholm, Sweden: IEEE), 3317–3322. doi:10.1109/ICRA.2016.7487505

Vega, A., Cintas, R., Manso, L. J., Bustos, P., and Núñez, P. (2020). "Socially-Accepted Path Planning for Robot Navigation Based on Social Interaction Spaces," in Fourth Iberian Robotics Conference: Advances in Robotics, Robot 2019, 644–655. doi:10.1007/978-3-030-36150-1_53

Vega, A., Manso, L. J., Cintas, R., and Núñez, P. (2019a). Planning Human-Robot Interaction for Social Navigation in Crowded Environments. *Adv. Phys. Agents* 855, 195–208. doi:10.1007/978-3-319-99885-5_14

Vega, A., Manso, L. J., Macharet, D. G., Bustos, P., and Núñez, P. (2019b). Socially Aware Robot Navigation System in Human-Populated and Interactive Environments Based on an Adaptive Spatial Density Function and Space Affordances. *Pattern Recognition Lett.* 118, 72–84. doi:10.1016/j.patrec.2018.07.015

Vega-Magro, A., Manso, L. J., Bustos, P., and Nunez, P. (2018). "A Flexible and Adaptive Spatial Density Model for Context-Aware Social Mapping: Towards a More Realistic Social Navigation," in 2018 15th International Conference on Control, Automation, Robotics and Vision, ICARCV, 1727–1732. doi:10.1109/ICARCV.2018.8581304

Vemula, A., Muelling, K., and Oh, J. (2018). "Social Attention: Modeling Attention in Human Crowds," in Proceedings - IEEE International Conference on Robotics and Automation (IEEE), 4601–4607. doi:10.1109/icra.2018.8460504

Walters, M. L., Dautenhahn, K., Te Boekhorst, R., Kheng Lee Koay, K. L., Kaouri, C., Woods, S., et al. (20052005). "The Influence of Subjects' Personality Traits on Personal Spatial Zones in a Human-Robot Interaction experiment," in ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 347–352. doi:10.1109/ROMAN.2005.1513803

Weina Ge, W., Collins, R. T., and Ruback, R. B. (2012). Vision-based Analysis of Small Groups in Pedestrian Crowds. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 1003–1016. doi:10.1109/tpami.2011.176

Xia, F., Zamir, A. R., He, Z., Sax, A., Malik, J., and Savarese, S. (2018). "Gibson Env: Real-World Perception for Embodied Agents," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (Salt Lake City: IEEE), 9068–9079. doi:10.1109/CVPR.2018.00945

Yan, Z., Duckett, T., and Bellotto, N. (2017). "Online Learning for Human Classification in 3D LiDAR-Based Tracking," in 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Basel, Switzerland: IEEE), 864–871. doi:10.1109/IROS.2017.8202247

Yang, C.-T., Zhang, T., Chen, L.-P., and Fu, L.-C. (2019). "Socially-aware Navigation of Omnidirectional mobile Robot with Extended Social Force Model in Multi-Human Environment," in Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics, Bari, Italy, 6-9 Oct. 2019 (Bari, Italy: IEEE), 1963–1968. doi:10.1109/smc.2019.8913844

Yang, F., and Peters, C. (2019). "Social-aware Navigation in Crowds with Static and Dynamic Groups," in 2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games), Vienna, Austria, 4-6 Sept. 2019 (Vienna, Austria: IEEE), 1–4. doi:10.1109/vs-games.2019.8864512

Yang, F., Yin, W., Bjorkman, M., and Peters, C. (2020). "Impact of Trajectory Generation Methods on Viewer Perception of Robot Approaching Group Behaviors," in 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 509–516. doi:10.1109/RO-MAN47096.2020.9223584

Yao, C., Liu, C., Liu, M., and Chen, Q. (2021). Navigation in Multi-Agent System with Side Preference Path Optimizer. *IEEE Access* 9, 113944–113953. doi:10.1109/ACCESS.2021.3104470

Yao, X., Zhang, J., and Oh, J. (2019). Following Social Groups: Socially Compliant Autonomous Navigation in Dense Crowds. ArXiv Prepr. ArXiv191112063, 2–5.

Yokoyama, N., Ha, S., and Batra, D. (2021). Success Weighted by Completion Time: A Dynamics-Aware Evaluation Criteria for Embodied Navigation. ArXiv210308022 Cs.

Yoon, H.-J., Widdowson, C., Marinho, T., Wang, R. F., and Hovakimyan, N. (2019). Socially Aware Path Planning for a Flying Robot in Close Proximity of Humans. *ACM Trans. Cyber-phys. Syst.* 3, 1–24. doi:10.1145/3341570

Young, J. E., Sung, J., Voida, A., Sharlin, E., Igarashi, T., Christensen, H. I., et al. (2011). Evaluating Human-Robot Interaction. *Int. J. Soc. Robotics* 3, 53–67. doi:10.1007/s12369-010-0081-8

Zhang, Y., Zhang, C.-H., and Shao, X. (2021). User Preference-Aware Navigation for mobile Robot in Domestic via Defined Virtual Area. *J. Netw. Comput. Appl.* 173, 102885. doi:10.1016/j.jnca.2020.102885

Zhong, S. W. J., Leong, L. W., and Ang, M. H., Jr. (2019). "Socially-acceptable Walking Parameters for Wheelchair Automation," in Proceedings of the Ieee 2019 9th International Conference on Cybernetics and Intelligent Systems (Cis) Robotics, Automation and Mechatronics (Ram) (Cis & Ram 2019), 193–197. doi:10.1109/cis-ram47153.2019.9095791

Zhou, Y., Wu, H., Cheng, H., Qi, K., Hu, K., Kang, C., et al. (2021). Social Graph Convolutional LSTM for Pedestrian Trajectory Prediction. *IET Intell. Transp. Syst.* 15, 396–405. doi:10.1049/itr2.12033

Zou, X., Sun, B., Zhao, D., Zhu, Z., Zhao, J., and He, Y. (2020). Multi-modal Pedestrian Trajectory Prediction for Edge Agents Based on Spatial-Temporal Graph. *IEEE Access* 8, 83321–83332. doi:10.1109/ACCESS.2020.2991435

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.