

Effects of Distinct Robot Navigation Strategies on Human Behavior in a Crowded Environment

Christoforos Mavrogiannis
Sibley School of Mechanical
and Aerospace Engineering
Cornell University
Ithaca, NY, USA
cm694@cornell.edu

Alena M. Hutchinson
Department of Computer Science
Cornell University
Ithaca, NY, USA
amh345@cornell.edu

John Macdonald
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA, USA
johnmacd@andrew.cmu.edu

Patrícia Alves-Oliveira
Instituto Universitário de Lisboa (ISCTE-IUL),
CIS-IUL, INESC-ID, Cornell University
Lisbon, Portugal
patricia_alves_oliveira@iscte-iul.pt

Ross A. Knepper
Department of Computer Science
Cornell University
Ithaca, NY, USA
rak@cs.cornell.edu

Abstract—State-of-the-art social robot navigation algorithms often lack a thorough experimental validation in human environments: simulated evaluations are often conducted under unrealistically strong assumptions that prohibit deployment in real world environments; experimental demonstrations that are limited in sample size do not provide adequate evidence regarding the user experience and the robot behavior; field studies may suffer from the noise imposed by uncontrollable factors from the environment; controlled lab experiments often fail to properly enforce challenging interaction settings. This paper contributes a first step towards addressing the outlined gaps in the literature. We present an original experiment, designed to test the implicit interaction between a mobile robot and a group of navigating human participants, under challenging settings in a controlled lab environment. We conducted a large-scale, within-subjects design study with 105 participants, exposed to three different conditions, corresponding to three distinct navigation strategies, executed by a telepresence robot (two autonomous, one teleoperated). We analyzed observed human and robot trajectories, under close interaction settings and participants' impressions regarding the robot's behavior. Key findings, extracted from a comparative statistical analysis include: (1) evidence that human acceleration is lower when navigating around an autonomous robot compared to a teleoperated one; (2) the lack of evidence to support the conventional expectation that teleoperation would be humans' preferred strategy. To the best of our knowledge, our study is unique in terms of goals, settings, thoroughness of evaluation and sample size.

Index Terms—Navigation; Motion Planning; Social Robotics.

I. INTRODUCTION

State-of-the-art autonomous navigation frameworks have been shown to achieve impressive benchmarks in simulation and to exhibit competent behaviors in experimental demonstrations, field studies, and lab experiments. However, their validation is often not sufficiently rigorous and in-depth. Simulated

C. Mavrogiannis and R. Knepper acknowledge support from the National Science Foundation, under Grants IIS-1526035 and IIS-1563705. P. Alves-Oliveira acknowledges support from the Fundação para a Ciência e a Tecnologia, under grants SFRH/BD/110223/2015 and UID/CEC/50021/2019.



Fig. 1: This study examines the performance of autonomous robot navigation algorithms in crowded pedestrian spaces.

evaluations are inevitably conducted under strong assumptions on the type of the environment, the context, and the type of behavior exhibited by other agents; thus addressing the *reality gap* problem is not a trivial extension. Experimental demonstrations contribute a significant step towards deployment to the real world but lack a significant sample size of repeated interaction with human users and thus statistical power. Large-scale field studies are an important step in the validation process of any robotic system as they may provide evidence of robust performance under challenging settings. Nonetheless, the noise induced by the frequently massive complexity of a

real-world environment may prohibit the extraction of concrete conclusions about the performance of the robot and the user experience. Lab studies may definitely isolate the system from external variables and enable rigorous testing of the desired conditions. However, designing an experiment that will isolate the desired nontrivial interaction between a target system and human participants is a not an easy task and it may often be observed that lab experiments with mobile robots do not test a challenging type of setting. Finally, ensuring the repeatability of the performance of an autonomous robotic system, exposed to close interaction with humans is also not a trivial task and often requires frequent maintenance and high costs. Thus, to approach interesting research questions without the complication of exhaustively testing the autonomy of the platform, a significant amount of research considers only Wizard of Oz experiments [33]. While the findings of such experiments are often of great significance regarding the human-robot interaction, they inevitably leave a gap in the validation of the autonomy itself.

A. Contributions

In this paper, we contribute a step towards addressing the outlined gap in the validation process of social navigation planning algorithms. We present an original experiment design, constructed to enforce naturally a series of challenging implicit interactions between a mobile robot and a group of human participants that navigate in a shared workspace in a controlled lab environment. A simple background scenario serves as a driving force towards interesting, nontrivial interactions but also as a way to cognitively load and distract human subjects from the goal of the experiment. We conducted an extensive, large-scale ($N = 105$), within-subjects user study in which we recorded 945 minutes of interaction between a mobile robot and human subjects (3 at a time). Participants are exposed to a set of three distinct robot navigation strategies, executed by a telepresence robot platform. We considered two autonomous navigation strategies and a teleoperated condition in which a human teleoperates the robot.

We collected human and robot trajectory data, recorded by an overhead motion capture system, and responses to a questionnaire designed to assess participants' impressions of the robot's intelligence, social compliance, and safety. We performed comparative statistical analyses on the collected datasets and report the results. Key findings, extracted by focusing on the close interactions (distance $< 1m$) between the robot and human subjects include the following: (1) human accelerations are significantly lower around an autonomous robot executing Social Momentum [29] than around a teleoperated one; (2) contrary to our expectations, we found no evidence to support the hypothesis that humans prefer the human-teleoperated navigation strategy—in fact humans did not distinguish between conditions in their ratings; (3) teleoperated motion that follows the same high-level rules as autonomy results in lower topological complexity [8] than autonomy, an observation potentially reflecting the more global character of human decision making for navigation tasks.

II. RELATED WORK

Social robot navigation constitutes a significant thrust of human-robot interaction research over the past few decades [24, 37, 6]. Particular emphasis has been placed on the design of autonomous socially aware navigation planning algorithms and on the study of the interaction between navigating robots and humans. Researchers have been inspired by the mechanisms underlying human navigation [45] and general human behavior in public spaces [11, 13]. This has led to the adoption of theory, models and methods from the field of cognitive science [30], psychology [44], sociology [45] and human-robot interaction [22] into the design and evaluation of proposed navigation frameworks. However, the complexity and cost of building and testing an autonomous robotic system often prohibits a systematic and thorough experimental validation of navigation frameworks. This section presents a classification of state-of-the-art approaches with respect to their type of validation.

A. Simulation Studies

Recent advances in the fields of graphics [15, 42, 17] and crowd dynamics [14, 16] were based on physics-inspired models of the interactions among multiple navigating agents: socially compliant and humanlike motion is generated as the result of multiple interacting potential fields, representing agents' objectives and intentions. This foundational idea has set the conceptual basis behind the design of a number of approaches in the field of social robot navigation.

Luber et al. [25] learn a set of dynamic navigation prototypes from a human trajectory dataset and use them for trajectory prediction and generation. They demonstrate the performance of their planner with respect to the efficiency and humanlikeness of generated paths on 182 scenarios of the same dataset and show how it outperforms a rule-based, proxemics-theory enforcing baseline. Vasquez et al. [43] learn a cost function to represent the dynamics of social navigation by training on a dataset extracted by teleoperating a robot in different real-world scenarios. They evaluate the ability of their learned model to reproduce trajectories of social compliance, modeled as a composite score of cost functions representing human comfort. Mavrogiannis et al. [28] learn a model that predicts the unfolding multi-agent trajectory topology [26, 27] in a crowded scene and use it to generate intent-expressive robot behaviors. Their approach is shown to simplify inference and decision making for co-present, heterogeneous agents in challenging simulated scenarios. An extension of their work, the Social Momentum planning framework [29] is shown to outperform two other baselines in terms of motion expressiveness, according to an online, video-based user study with 180 human participants. Finally, Bera et al. [5] make use of concepts from Personality Trait theory to classify the behavior of pedestrians towards informing their motion models and a robot's path planning. Simulation results demonstrate improved trajectory prediction and more socially compliant on a number of human datasets.

B. Experimental Demonstrations

A number of works have presented important experimental demonstrations in human environments to validate their approaches. Althaus et al. [2] focus on the problem of social engagement. They build a robot designed to approach humans and engage in a conversation with them. They present control laws for approaching a person and maintaining a socially acceptable distance. A recorded experiment with three participants demonstrates the efficacy of their approach. Sisbot et al. [35] presents a cost-based planner that considers a set of social heuristics at the planning stage to generate motion that is visible and safe around humans. A series of documented interactions between the robot and a human in a lab environment demonstrate the capabilities of the framework. Park et al. [32] build an automated wheelchair and design a model-predictive control law for smooth motion generation in crowded environments. Their approach treats humans as dynamic obstacles and focuses on avoiding them smoothly. They test their framework in an indoor environment and report a set of successful collision avoidance processes under crowded settings. Kretschmar et al. [23] employ an inverse reinforcement learning approach to learn the social components of human navigation by training on an hour-long lab dataset of four navigating humans and on a public dataset on a crowded scene. They deploy the model on a robotic wheelchair that is able to navigate socially next to navigating humans in a narrow hallway. Finally, Chen et al. [7] present a deep reinforcement learning approach to learn social norms (passing from the right-hand side and overtaking on the left) from a synthetic, simulated dataset. They report an experimental demo, run in a large, crowded academic building.

C. Experimental Studies

A significant body of work has employed field studies in crowded environments such as museums, malls or academic hallways. Thrun et al. [38] present a tour-guide robot equipped with a set of probabilistic algorithms for mapping, localization, people-tracking, and planning. The robot interacted successfully with thousands of visitors for two weeks in a busy museum. The authors present a comprehensive report of the robot's log and a classification of observed types of interaction between the robot and visitors. Bennewitz et al. [4] cluster a dataset of observed human trajectories into a set of classes and use it for on-line prediction on a robot. A series of 10 experiments indicates increased time-efficiency resulting from their approach, compared to a linear prediction baseline. Pacchierotti et al. [31] implement and test a proxemics-based control framework on an autonomous robot through a user study, conducted in a corridor. A total of 10 participants were exposed to three different conditions corresponding to the robot passing next to them with a different lateral distance each time. Users' ratings showed that humans felt uncomfortable when the robot was closer to them. Foka and Trahanias [9] present a probabilistic algorithm that makes predictions about future human paths to plan collision-free motion. They report logs and performance aspects upon running the robot

for 70 hours in an indoor academic building. Kirby et al. [21] present a constrained optimization-based algorithm that incorporates a series of social conventions, such as passing-side conventions and respect of humans' personal space into the robot's decision making. A user study involving 27 human subjects navigating alongside a robot in an academic hallway demonstrated evidence that humans interpret the robot's behavior as socially appropriate [20]. Shiomi et al. [34] present a planner, based on the social force model [14] for generating humanlike collision avoidance navigation behaviors. A 4-hour field study in a shopping mall demonstrated that the proposed approach achieves safer and more comfortable interaction than a baseline. Trautman et al. [40] present a navigation framework that explicitly incorporates the assumption of human cooperation into their learned trajectory prediction model to enable a robot to navigate among dense human crowds. They report the performance of a real robot in terms of safety and efficiency in a large-scale field study (488 runs), conducted in a crowded cafeteria. Kato et al. [18] learn a model of human intent inference to generate social approaching navigation behaviors. They test their approach on a humanlike robot employee in a crowded mall and record interactions with 130 people, suggesting that a compromise between proactive and passive approaching behavior is preferred by humans. Kim and Pineau [19] learn a model of socially compliant robot motion from human demonstrations and robot teleoperations in crowded environments. They test their approach on a robotic wheelchair in a crowded hallway and report humanlike and efficient performance in 10 field runs. Truong and Ngo [41] fuse elements of the social force model [14] and the Reciprocal Velocity Obstacle model [42] to generate socially aware robot motion in crowded scenes. Examples from experiments, conducted in an office environment demonstrate smooth operation against static or moving obstacles.

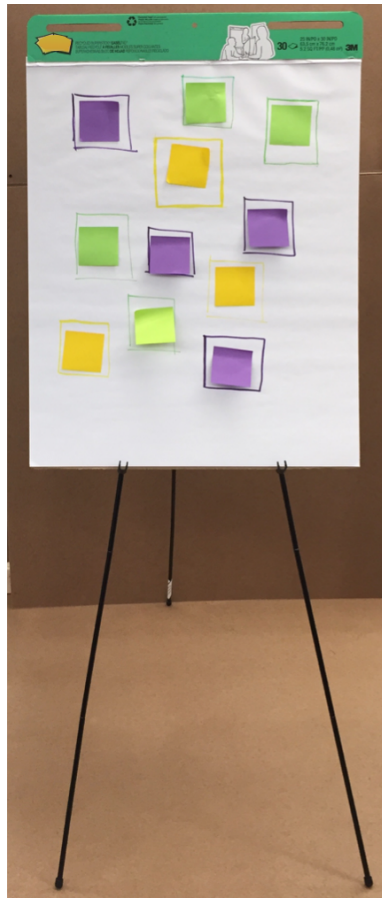
III. USER STUDY

In this paper we present an IRB-approved (approval code: 1805008009) lab study, focused on the evaluation of a set of distinct robot navigation algorithms with respect to social compliance. The lab environment allows us to have significant control over variables that can interfere with the experimental setting. We leverage this level of control to enforce challenging navigation behaviors in a natural fashion through the design of an original experiment scenario and task.

We enforce a setting of implicit, nonverbal social engagement among agents, similar to the type of interaction among walking pedestrians so that we can study phenomena involving collaborative collision avoidance processes, as observed by Wolfinger [45]. Furthermore, we construct a moderately crowded scene that balances close interactions with space for the robot to showcase its distinct navigation strategies (see Fig. 1). We also ensure the emergence of nontrivial interactions, involving challenging collision-avoidance maneuvers between participants and the robot through the definition of rules. Moreover, we motivate natural walking behaviors by not disclosing the real purpose of the study until the debriefing



(a) Robot.



(b) Easel.



(c) Construction helmets with trackers, sticky notes, and markers.
Fig. 2: Study apparatus: Beam Pro robot [1] (Fig. 2a); an easel, representing a machine with sticky notes, representing completed maintenance tasks (Fig. 2b); tracking helmets, sticky notes and markers distributed to participants (Fig. 2c).

process and by increasing participants' cognitive load through the background scenario and task. Finally, we consider three conditions and keep the total duration less than thirty minutes to facilitate recruiting and minimize potential effects resulting from participants' fatigue.

A. General Experiment Procedure

Our study is organized into a set of experiment sessions. In each session, three different human subjects participate

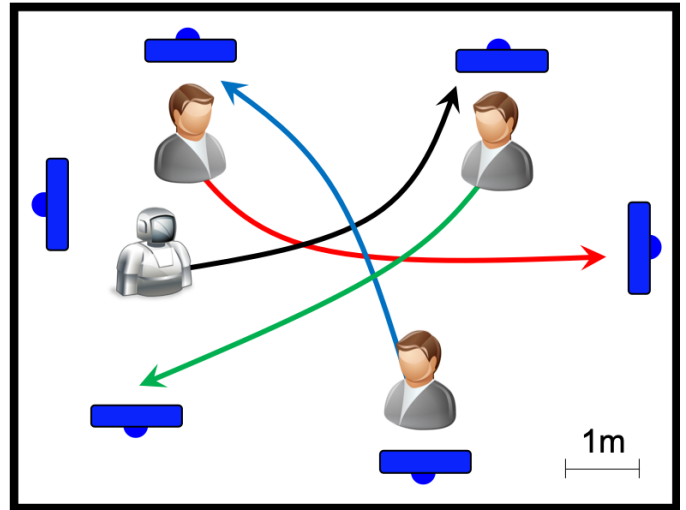


Fig. 3: Top view of the workspace along with example human and robot trajectories, corresponding to transitions between easels (blue objects).

in a set of three experiment trials. Before the first trial, participants are asked to give written consent to confirm their participation and optionally to be video recorded. A member of our research team delivers the instructions and answers questions. During each trial, participants repeatedly visit a set of stations inside a rectangular workspace of area $16m^2$ (see Fig. 3), driven by a fictional scenario. In parallel, a mobile robot (a Suitable Technologies Beam Pro, equipped with a quad core i7 processor laptop from 2017), shown in Fig. 2a, also moves between the stations. During each trial, the human and robot trajectories are tracked and recorded through an overhead motion capture system of six high-accuracy ($< 1\text{ mm}$), high-fidelity (frequency 180 Hz) cameras and videotaped if participants gave consent. Real-time tracking was enabled through the use of construction helmets (see Fig. 2a, Fig. 2c) with reflective markers. After each trial, participants are asked to fill in a questionnaire, containing questions about their impressions from their interactions with the robot. At the end, participants are asked to provide basic demographic data and information regarding their prior experience with user studies and robotics technology. Participants are then debriefed, compensated and dismissed.

B. Background Scenario and Task Description

Participants are asked to imagine that they are workers in a factory (the factory setting helps justify the tracking helmets) and the robot is a supervisor. The factory environment (lab workspace) contains six machines, represented as easels (see Fig. 2b), spread around the workspace, as shown in Fig. 3. Each worker is given a distinctly colored marker and a contrasting, distinctly colored set of sticky notes (see Fig. 2c). The duty of a worker is to **perform** maintenance tasks to machines and **assign** tasks for other workers to perform. Assigning a task is done by drawing a square on the pad of an easel, whereas performing a task is done by posting a sticky note inside a square drawn on an easel pad (see Fig. 2b). Participants are asked to perform only tasks represented with

squares of color that matches the color of their sticky notes.

C. Trial Description

Before the start of each trial, participants are randomly positioned next to different machines and the robot is placed in the middle of the workspace, as shown in Fig. 3. A trial is organized into a set of maintenance cycles, initiated by a gong sound, played by the robot. Each time the gong is played, participants are instructed to leave their machines towards a non-adjacent machine of their choosing. Each time participants reach a new machine, they are instructed to perform up to one pre-assigned task (if one exists) and assign a new task. At the same time, the robot is navigating in the workspace by following the same rules of transitioning between stations, i.e. it only moves to a randomly picked, non-adjacent machine when the gong sound is played. For synchronization purposes, the gong sound is played when the robot is ready to move towards its next machine. Each trial lasts exactly three minutes, during which an ambient factory sound track is played.

D. Conditions

All participants were exposed to the same three conditions (within-subjects design), each corresponding to a different navigation strategy, executed by the robot. To account for potential ordering effects (i.e., due to fatigue, frustration, learning), the condition order was methodically varied and approximately equally spread across all sessions. The selected set of navigation strategies consists of *Optimal Reciprocal Collision Avoidance* (ORCA) [42], *Social Momentum* (SM) [29]) and teleoperation (TE). These strategies were mainly selected due to the diversity of decision making principles that they represent, i.e., ORCA is designed to be optimal, SM is inherently intention-aware; TE is designed to appear humanlike. Additional reasons that influenced our selection included: (1) the fact that ORCA constitutes a common benchmark and work of reference for multi-agent simulations (e.g. [7, 10, 23, 29]); (2) the existence of an open source, optimized C++ implementation of ORCA; (3) the ease of implementation of SM; (4) the widespread use of telepresence robot platforms through teleoperation via their navigation interfaces. The complexity of a real-world pedestrian environment would pose a significant challenge to any of these navigation planners. However, we believe that an extensive and comparative evaluation of planners with distinct philosophies provides us with significant insights and experience for the design of the next generation of social navigation planning algorithms. The following paragraphs provide short descriptions of the mechanisms underlying the selected navigation strategies.

Optimal Reciprocal Collision Avoidance (commonly referred to as ORCA; in the results section of the paper we will be using the codename OR for brevity) [42] is a decentralized navigation planning framework for the generation of smooth, collision-free, natural-looking simulations of multi-agent scenarios. It is an optimization-based approach that determines the velocity of minimal divergence from an agent's desired velocity that is guaranteed to be collision-free for a desired

time horizon. This approach makes local collision avoidance considerations by incorporating a model of intentions, based on agents' current velocities. It operates however, under the assumption that other agents also run the same decision making mechanism to guarantee safe and smooth behaviors.

Social Momentum (SM) [29] is a decentralized, cost-based, navigation planner, designed to generate legible robot motion in multi-agent environments. The cost is a weighted sum of two functions: one representing efficiency and one representing social compliance. At planning time, the robot selects and executes the action that contributes the best compromise between the two costs. This policy results in consistent progress towards an agent's destination while taking into consideration the collision-avoidance intentions and preferences of other agents. Unlike ORCA, SM does not explicitly assume that others run the same policy; instead, it focuses on reading the intentions of others and incorporates this knowledge into its motion planning process.

The *Teleoperation* strategy (TE) was implemented through the official navigation interface provided by the manufacturer [1], using the arrow keys on a standard laptop keyboard. This interface contains two live streams of video, providing the teleoperator with real-time video streams of a forward, wide-angle field of view (top) and a floor view (bottom). Navigation commands may be executed through a simple keyboard's arrow keys (or with a mouse). Selected commands are demonstrated as projected future trajectories on the video streams, providing visual feedback to the user. The teleoperation condition was executed by the same member of our research team across all sessions, from a remote location (outside of the lab). The teleoperator had significant prior experience of the navigation interface for several years. Before collecting data for our final dataset, we completed a total of 7 rounds of pilot sessions under different variants of the final study setup. Thus by the time we officially started the study, the teleoperator had reached a skill level that qualitatively appeared to be appropriate for the needs of the condition. Although it is hard to precisely quantify the operator's skill level, his experience was in the order of several hours prior to the start of the study and thus we do not believe that his performance evolved over the course of the study as a result of learning.

E. Hypotheses

Upon experimenting with the three navigation strategies considered (simulations conducted with SM and ORCA, personal teleoperated teleconference sessions with the Beam), we observed very different patterns of decision making. These patterns were interpreted as the result of the different design principles and objectives behind each framework: ORCA was developed to produce efficient, realistic simulations of virtual multi-agent scenarios; SM was designed to generate legible robot motion in dynamic multi-agent environments; TE was based on a navigation interface [1], specifically designed to allow non-expert users to control a robot intuitively. To the best of our knowledge, these strategies have never been tested

against each other under challenging, multi-agent, experimental settings. It was unclear how close interaction between the robot and different human participants would affect the motion generated by the different strategies. Furthermore, it was uncertain how humans would react to different behaviors exhibited by the robot and how this interaction would affect overall performance for both humans and the robot. Using the dataset generated by our study, we explore these questions by examining the validity of the following hypotheses:

(H1) - Robot Performance: In close interactions with humans: (a) ORCA generates the most geometrically efficient paths; (b) SM generates the jerkiest paths; (c) TE generates the most energy-efficient paths.

(H2) - Human performance: Humans navigating in close proximity with the robot: (a) follow the least jerky paths when the robot runs SM; (b) spend the least energy when the robot runs TE; (c) spend the most energy when the robot runs OR.

(H3) - Group performance: Global group (human and robot) behavior under SM results in trajectories of lower Topological Complexity than the other two conditions.

(H4) - Human Impressions: Participants consider the behaviors generated by TE as more socially compliant, intelligent and safe than the rest of the strategies.

IV. ANALYSIS

We conducted 35 experiment sessions, in which a total of 105 human subjects were exposed to all three conditions. Subjects were recruited from a university population (Cornell University), through a centralized, university-run subject-recruitment website and also through fliers posted across campus. The subjects (59 female, 45 male, 1 unidentified) were 21.45 years old on average ($SD = 3.19$ years) with their age ranging from 18 to 33 years. About half of them (57) had prior experience of user study participation and they rated their familiarity with robotics technology with an average of 2.47 ($SD = 1.27$) on a 5-point Likert scale.

We collected a dataset comprising the trajectories of all 105 participants and the robot across all trials. Focusing on dynamic interactions of close proximity, we split this raw dataset into two datasets of trajectory segments: (a) a dataset comprising 1033 robot trajectory segments of close interaction with humans (minimum distance $d < 1m$) and (b) a dataset comprising 1566 human trajectory segments of close interaction with the robot (also, of minimum distance $d < 1m$). We analyze the trajectory dataset using a set of trajectory quality measures from relevant literature [23, 12, 29], computed over fixed timestep intervals (100 timesteps, totaling 0.2 seconds). In particular, we computed: (1) the average *Acceleration* per segment, a ; (2) the average *Energy* per segment, E , where energy is defined as the integral of the squared velocity of an agent throughout its trajectory; (3) the minimum *Distance* between the robot and any other humans per segment, d ; (4) *Path Irregularity* per segment, PI , measuring the total amount of unnecessary rotation (angle between an agent's heading and direction to goal) that an agent exhibits per unit path length [12]; (5) *Path Efficiency*, \mathcal{E} , defined as the ratio of the distance

between the endpoints of a segment over the length of the path that the agent actually followed; (6) time spent per unit path length over a segment, τ ; (7) *Topological Complexity*, TC [8, 29], defined as the amount of entanglement among agents' trajectories throughout a trial (the Braidlab software package [36] was used for these computations).

We also collected a dataset comprising the responses of all 105 participants to a questionnaire, containing Likert-scale style questions, based on the instrument of Bartneck et al. [3] and short response questions.

TABLE I: Effect of Navigation Strategy on Robot Behavior

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
a	56.09	28.05	2	65.12	58.94	< 0.0001
E	0.7083	0.3541	2	1015	440.1	< 0.0001
\mathcal{E}	0.05796	0.02898	2	999.1	4.825	0.008213
PI	454.4	227.2	2	1012	355.3	< 0.0001
τ	116.5	58.27	2	1016	1056	< 0.0001

A. Effect of Navigation Strategy on Robot Behavior

We model the effect of condition (ORCA, SM, TE) on each one of the trajectory quality measures considered. We use linear mixed-effects regression models, to account for both fixed effects resulting from the conditions but also for random effects resulting from the session and the trial (expected means with confidence intervals are depicted in Fig. 4).

One-way ANOVA performed on the models demonstrates a significant effect of the condition on all robot trajectory quality criteria at the $p < 0.05$ level (see table I for the test statistics and Fig. 4 for the expected means and confidence intervals for all criteria) and thus, we find that (H1) is confirmed. More specifically, it can be observed that ORCA generates the smoothest motion among all strategies (lowest acceleration, lowest path irregularity, lowest time), which confirms (H1a). This trend was expected as ORCA selects actions that minimize divergence from an agent's direction to goal and desired speed to ensure collision avoidance for a desired time window. This results in a smoother speed profile than other conditions. SM on the other hand, prioritizes intent-expressiveness by exaggerating its motion to indicate an intended passing-side intention; this results in higher acceleration (due to rotation) and path irregularity, which confirms (H1b). Finally, TE is the most energy-efficient — which confirms (H1c) — but also the least time-efficient of all strategies. These findings could mainly be attributed to the defensive driving style of the teleoperator and the navigation through arrow keys.

B. Effect of Navigation Strategy on Human Behavior

Similarly to robot trajectory, we model the dependency of the human trajectory quality measures to the condition with linear mixed-effects models, accounting also for random effects of session, trial and helmet per trial. Fig. 5 depicts the expected means and confidence intervals for the human trajectory quality measures, whereas table II contains statistics extracted upon performing ANOVA on the models at the $p < 0.05$ significance level.

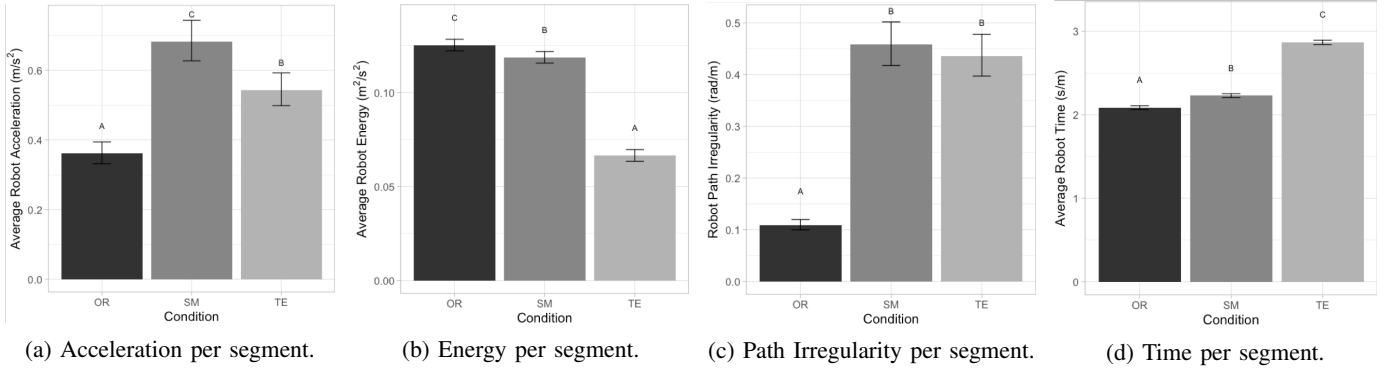


Fig. 4: Expected means and confidence intervals for robot trajectory quality criteria. Quantities labeled with distinct letters (A, B, C) come from significantly different distributions (Tukey’s HSD test, $p < 0.05$).

TABLE II: Effect of Navigation Strategy on Human Behavior

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
<i>a</i>	1.415	0.7073	2	250.4	3.888	0.02173
<i>d</i>	0.1075	0.05377	2	231.5	0.5872	0.5567
<i>E</i>	0.112	0.05599	2	253.3	3.449	0.03326
<i>E</i>	0.02977	0.01489	2	68.46	1.959	0.1488
<i>PI</i>	0.5394	0.2697	2	249.4	3.286	0.03904
τ	0.08277	0.04139	2	252.7	2.145	0.1192

Overall, we find that (H2) is confirmed. In particular, we see that humans exposed to the SM condition followed smoother trajectories, of lower acceleration (Fig. 5a) and path irregularity (Fig. 5c) than humans exposed to either ORCA or TE, which confirms (H2a). This was in line with our expectations: SM’s intention-aware navigation strategy adapts the robot’s behavior to the preferences of humans, thus facilitating human inference and decision making. Further, it was observed that humans spend the least energy when exposed to TE, which confirms (H2b). We attribute this finding to the perceived humanlikeness of the motion generated by a teleoperated robot: the embodiment of human decision making on a robot platform features humanlike traits that potentially enable a higher level of human comfort. Finally, humans spend the most energy around OR (see Fig. 5b), which confirms (H2c). This could be perceived as an result of ORCA’s more predictable motion (minimal divergence from desired direction). Higher predictability potentially results in higher confidence for participants, which allows them to move faster and thus spend more energy.

C. Effect of Navigation Strategy on Group Behavior

We model the effect of condition on the Topological Complexity of the group trajectory (the set of all agents’ trajectories) over a trial, using a linear mixed-effects model (accounting for random effects of session, trial and helmet per trial). Overall, we find that (H3) is rejected. ANOVA performed on the model uncovered a significant variance among conditions ($F(2, 67.71) = 8.075, p = 0.000716$, see table III, Fig. 5d). In particular, it was found that the Topological Complexity of trajectories, generated by groups exposed to TE was significantly lower than both SM and OR. Global group behavior generated in the presence of autonomy was significantly more complex, despite the fact that the human teleoperator was following the same rule for transitioning

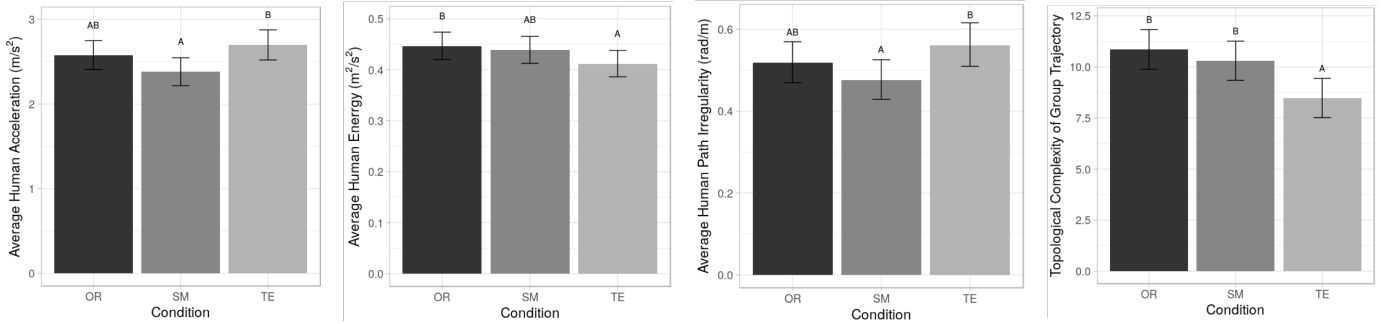
between machines (random selection of any non-adjacent machine). In other words, autonomous strategies resulted in more intense mixing among all four agents. This could be attributed to the mechanisms underlying human navigation, as the decision making computations under TE were done by the human teleoperator. Lower *TC* represents trajectory entanglement which intuitively corresponds to behaviors of passing *around* as opposed to passing *through* others. Thus, this trend could be attributed to the tendency of the human teleoperator to avoid collisions more *globally*, by avoiding any type of encounter with other participants whereas the robot was employing a more local collision avoidance mechanism by sequentially responding to any challenging encounters. This finding is perhaps unsurprising since both autonomous algorithms considered explicitly favor the avoidance of closer collisions over further ones.

TABLE III: Effect of Navigation Strategy on Group Behavior

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
<i>TC</i>	107.3	53.66	2	67.71	8.075	0.000716

D. Effect of Navigation Strategy on Human Impressions

We model the effect of condition to each of the Likert scale questions considered, using a Linear Mixed Effects Regression Model. Table IV contains a list of the questions that were posed to participants (as 5-point Likert scales, with 1 denoting a negative response and 5 denoting a positive response), grouped into three different classes: (a) one referring to the robot’s behavior (orange rows); (b) one referring to participants’ emotional states during the experiment (yellow rows); (c) one referring to participants’ expectations about the future presence of the robot (blue rows). The table also contains the statistics of one-way ANOVA tests, performed to participants’ responses to each question. Significant variance was observed in the responses to the question about the robot’s intelligence ($F(2, 269.73) = 3.115, p = 0.0460$) and it was found that participants rated the intelligence of TE as slightly higher ($M = 3.29, SE = 0.11$) than both SM ($M = 3.01, SE = 0.11$) and OR ($M = 3.04, SE = 0.11$). This trend also suggests a potential perception of the humanlikeness of TE from the perspective of the participants, which appears to be in line with the fact that they spent significantly less energy around TE. However, this trend is not reflected in the



(a) Acceleration per segment.

(b) Energy per segment.

(c) Path Irregularity per segment.

(d) Top. Complexity per trial.

Fig. 5: Expected means and confidence intervals for human (Fig. 5a, Fig. 5b, Fig. 5c) and group trajectory quality criteria (Fig. 5d). Quantities labeled with distinct letters (A, B) are significantly different (Tukey's HSD test, $p < 0.05$).

responses to the rest of the questions. Therefore, we cannot conclusively confirm or reject (H4). It might be the case that the quantitative differences among conditions in terms of the quality criteria were below the precision of human perception.

TABLE IV: Effect of Nav. Strategy on Human Impressions

Rating	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
Competent	3.249	1.625	2	269.1	1.875	0.1553
Responsible	0.546	0.2732	2	269.7	0.3047	0.7376
Predictable	0.8769	0.4384	2	270.6	0.3760	0.6870
Compliant	3.599	1.800	2	269.7	2.279	0.1044
Sensible	1.687	0.8435	2	269.7	0.9756	0.3783
Friendly	2.346	1.173	2	269.6	1.299	0.2745
Safe	4.5409	2.270	2	270.0	1.793	0.1684
Pleasant	1.8920	0.9460	2	269.5	1.149	0.3184
Polite	0.9942	0.4971	2	269.8	0.7171	0.4891
Coordinated	2.971	1.485	2	269.4	1.425	0.2423
Intelligent	4.840	2.420	2	269.7	3.115	0.0460*
Trustworthy	0.5773	0.2887	2	269.2	0.4823	0.6179
Socially aware	4.507	2.254	2	268.6	2.131	0.1207
Discreet	4.742	2.371	2	270.2	2.238	0.1087
Relaxed	2.603	1.301	2	269.1	1.210	0.2998
Calm	2.405	1.202	2	269.7	1.123	0.3268
Tranquil	0.1069	0.0535	2	269.4	0.0575	0.9441
Noticeable	3.139	1.570	2	270.3	1.341	0.2633
Predict (future)	0.4635	0.2318	2	270.5	0.1937	0.8241
Bump (future)	7.572	3.786	2	270.2	2.682	0.0702

V. DISCUSSION

We presented a within-subjects user study design for the experimental evaluation of mobile robot navigation strategies in a controlled lab environment. Our experiments involved the navigation of a mobile robot in a workspace shared with three human participants, under challenging settings of implicit interaction, emulating aspects of pedestrian navigation. We conducted a total of 35 experiment sessions in which 105 human participants were exposed to the same set of conditions corresponding to three different navigation strategies executed by the robot. We analyzed the collected dataset through the use of objective measures (trajectory analysis) and subjective measures (questionnaires asking for ratings of participants' impressions of robot's intelligence, safety and personality). We found statistical evidence that humans follow less jerky and irregular paths when navigating around one autonomous navigation condition [29] than around a teleoperated robot. Furthermore, contrary to our expectations, humans did not discriminate between conditions, according to their responses to our questionnaire. Finally, we presented evidence that human decision making, as captured in the teleoperated condition, had

a more global character than the autonomous strategies. We plan to investigate this finding further in future work.

A. Limitations

Our study encompasses some limitations generally inherent to any HRI study, and some specific to our scenario. First, a controlled lab environment cannot emulate the complexity of a real-world pedestrian environment and no background scenario or task could give rise to perfectly natural human walking behaviors. Furthermore, humans lack models of interaction with robotic technology, which inevitably affects their behavior around a mobile robot. Even the robot's appearance, structure and dynamics could attract attention and distract participants from the task. Moreover, the selection of the navigation strategies inevitably impacts the generalizability of the results. Either of the autonomy conditions could struggle with erratic human behavior and specifically with human motion that is suboptimal with respect to intent and flexibility [39] whereas the teleoperator's performance may vary across individuals, experience, skill level, driving style, etc. Finally, the sample of participants, mostly coming from the undergraduate population of a university introduces another confound.

B. Broader Impact

Despite its many limitations, this study is unique in terms of its goals, settings, thoroughness of evaluation and sample size. As stated in the introduction, this study was motivated by an observed gap in the literature: we believe that the validation of social navigation algorithms requires a more thorough process. The stage of a controlled lab evaluation is an indispensable part of the validation process and should not be discounted before deploying a robot to the field. The field of social robot navigation could benefit greatly from extensive in-lab validation of additional algorithms, under various interaction settings. We hope that this study will constitute a paradigm for such future studies in terms of its design and scope and a reference for informing the design of future algorithms, within the field of navigation and beyond.

ACKNOWLEDGMENT

The authors would like to thank Erika Louise Mudrak, from the Statistical Consulting Unit at Cornell University for her valuable advice on the analysis of the collected datasets.

REFERENCES

- [1] Suitable technologies, Inc. <https://suitabletech.com/>. Accessed: 2018-09-13.
- [2] P. Althaus, H. Ishiguro, T. Kanda, T. Miyashita, and H. I. Christensen. Navigation for human-robot interaction tasks. In *Proceedings of the 2004 IEEE International Conference on Robotics and Automation (ICRA '04)*, volume 2, pages 1894–1900 Vol.2, 2004.
- [3] Christoph Bartneck, Elizabeth Croft, and Dana Kulic. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, 2009.
- [4] Maren Bennewitz, Wolfram Burgard, Grzegorz Cielniak, and Sebastian Thrun. Learning motion patterns of people for compliant robot motion. *International Journal of Robotics Research*, 24:31–48, 2005.
- [5] A. Bera, T. Randhavane, R. Prinja, and D. Manocha. Sociosense: Robot navigation amongst pedestrians with social and psychological constraints. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '17)*, pages 7018–7025, 2017.
- [6] Konstantinos Charalampous, Ioannis Kostavelis, and Antonios Gasteratos. Recent trends in social aware robot navigation: A survey. *Robotics and Autonomous Systems*, 93:85 – 104, 2017.
- [7] Y. F. Chen, M. Everett, M. Liu, and J. P. How. Socially aware motion planning with deep reinforcement learning. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '17)*, pages 1343–1350, 2017.
- [8] Ivan Dynnikov and Bert Wiest. On the complexity of braids. *Journal of the European Mathematical Society*, 009(4):801–840, 2007.
- [9] Amalia F. Foka and Panos E. Trahanias. Probabilistic autonomous robot navigation in dynamic environments with human motion prediction. *International Journal of Social Robotics*, 2(1):79–94, Mar 2010.
- [10] Julio Godoy, Ioannis Karamouzas, Stephen J. Guy, and Maria Gini. Moving in a crowd: Safe and efficient navigation among heterogeneous agents. pages 294–300, 2016.
- [11] Erving Goffman. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. Free Press, 1966.
- [12] J. Guzzi, A. Giusti, L. M. Gambardella, G. Theraulaz, and G. A. Di Caro. Human-friendly robot navigation in dynamic environments. In *Proceedings of the 2013 IEEE International Conference on Robotics and Automation (ICRA '13)*, pages 423–430, 2013.
- [13] E.T. Hall. *The Hidden Dimension*. Anchor Books, 1990.
- [14] Dirk Helbing and Péter Molnár. Social force model for pedestrian dynamics. *Physical Review E*, 51:4282–4286, 1995.
- [15] Peter Henry, Christian Vollmer, Brian Ferris, and Dieter Fox. Learning to navigate through crowded environments. In *Proceedings of the 2010 IEEE International Conference on Robotics and Automation (ICRA '10)*, pages 981–986, 2010.
- [16] Serge Hoogendoorn and Piet HL Bovy. Simulation of pedestrian flows by optimal control and differential games. *Optimal Control Applications and Methods*, 24(3):153–172, 2003.
- [17] Ioannis Karamouzas, Brian Skinner, and Stephen J. Guy. Universal power law governing pedestrian interactions. *Physical Review Letters*, 113:238701, 2014.
- [18] Yusuke Kato, Takayuki Kanda, and Hiroshi Ishiguro. May I help you?: Design of human-like polite approaching behavior. In *Proceedings of the 2015 ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*, pages 35–42, 2015.
- [19] Beomjoon Kim and Joelle Pineau. Socially adaptive path planning in human environments using inverse reinforcement learning. *International Journal of Social Robotics*, 8(1):51–66, 2016.
- [20] Rachel Kirby. *Social Robot Navigation*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, May 2010.
- [21] Rachel Kirby, Reid Simmons, and Jodi Forlizzi. Companion: A constraint-optimizing method for person-acceptable navigation. In *Proceedings of the 2009 IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '09)*, pages 607–612, 2009.
- [22] Ross A. Knepper, Christoforos I. Mavrogiannis, Julia Proft, and Claire Liang. Implicit communication in a joint action. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*, pages 283–292, 2017.
- [23] Henrik Kretzschmar, Markus Spies, Christoph Sprunk, and Wolfram Burgard. Socially compliant mobile robot navigation via inverse reinforcement learning. *The International Journal of Robotics Research*, 35(11):1289–1307, 2016.
- [24] Thibault Kruse, Amit Kumar Pandey, Rachid Alami, and Alexandra Kirsch. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems*, 61(12):1726–1743, 2013.
- [25] M. Luber, L. Spinello, J. Silva, and K. O. Arras. Socially-aware robot navigation: A learning approach. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '12)*, pages 902–907, 2012.
- [26] Christoforos I. Mavrogiannis and Ross A. Knepper. Decentralized multi-agent navigation planning with braids. In *Proceedings of the 2016 International Workshop on the Algorithmic Foundations of Robotics (WAFR '16)*, 2016.
- [27] Christoforos I. Mavrogiannis and Ross A. Knepper. Multi-agent path topology in support of socially competent navigation planning. *The International Journal of Robotics Research*, 2018.

- [28] Christoforos I. Mavrogiannis, Valts Blukis, and Ross A. Knepper. Socially competent navigation planning by deep learning of multi-agent path topologies. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '17)*, pages 6817–6824, 2017.
- [29] Christoforos I. Mavrogiannis, Wil B. Thomason, and Ross A. Knepper. Social momentum: A framework for legible navigation in dynamic multi-agent environments. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*, pages 361–369, 2018.
- [30] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, 108(17):6884–6888, 2011.
- [31] E. Pacchierotti, H. I. Christensen, and P. Jensfelt. Evaluation of passing distance for social robots. In *Proceedings of the 2006 IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '06)*, pages 315–320, 2006.
- [32] Jong Jin Park, Collin Johnson, and Benjamin Kuipers. Robot navigation with model predictive equilibrium point control. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '12)*, pages 4945–4952, 2012.
- [33] Laurel D Riek. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1):119–136, 2012.
- [34] Masahiro Shiomi, Francesco Zanlungo, Kotaro Hayashi, and Takayuki Kanda. Towards a socially acceptable collision avoidance for a mobile robot navigating among pedestrians using a pedestrian model. *International Journal of Social Robotics*, 6(3):443–455, 2014.
- [35] Emrah Akin Sisbot, Luis Felipe Marin-Urias, Rachid Alami, and Thierry Siméon. A human aware mobile robot motion planner. *IEEE Transactions on Robotics*, 23(5):874–883, 2007.
- [36] Jean-Luc Thiffeault and Marko Budišić. Braidlab: A software package for braids and loops, 2013–2017. URL <http://arXiv.org/abs/1410.0849>. Version 3.2.3.
- [37] Andrea Thomaz, Guy Hoffman, and Maya Cakmak. Computational human-robot interaction. *Foundations and Trends in Robotics*, 4(2-3):105–223, 2016.
- [38] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hhnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva. *The International Journal of Robotics Research*, 19(11):972–999, 2000.
- [39] Pete Trautman. Sparse interacting gaussian processes: Efficiency and optimality theorems of autonomous crowd navigation. In *Proceedings of the 2017 IEEE Conference on Decision and Control (CDC '17)*, pages 327–334, 2017.
- [40] Peter Trautman, Jeremy Ma, Richard M. Murray, and Andreas Krause. Robot navigation in dense human crowds: Statistical models and experimental studies of human-robot cooperation. *International Journal of Robotics Research*, 34(3):335–356, 2015.
- [41] X. T. Truong and T. D. Ngo. Toward socially aware robot navigation in dynamic and crowded environments: A proactive social motion model. *IEEE Transactions on Automation Science and Engineering*, 14(4):1743–1760, 2017. ISSN 1545-5955.
- [42] Jur van den Berg, Stephen J. Guy, Ming C. Lin, and Dinesh Manocha. Reciprocal n -body collision avoidance. In *Proceedings of the International Symposium on Robotics Research (ISRR '09)*, pages 3–19, 2009.
- [43] D. Vasquez, B. Okal, and K. O. Arras. Inverse reinforcement learning algorithms and features for robot navigation in crowds: An experimental comparison. In *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '14)*, pages 1341–1346, 2014.
- [44] William H. Warren. The dynamics of perception and action. *Psychological Review*, 113(2):358–389, 2006.
- [45] Nicholas H. Wolfinger. Passing Moments: Some Social Dynamics of Pedestrian Interaction. *Journal of Contemporary Ethnography*, 24(3):323–340, 1995.