Robots Asking for Directions – The Willingness of Passers-by to Support Robots

Astrid Weiss, Judith Igelsböck, Manfred Tscheligi
ICT&S Center
University of Salzburg
Salzburg, Austria
Email: firstname.lastname@sbg.ac.at

Andrea Bauer, Kolja Kühnlenz, Dirk Wollherr, Martin Buss
Institute of Automatic Control Engineering
Technische Universität München
Munich, Germany
Email: ab, koku, dw, mb@tum.de

***Abstract*—This paper reports about a human-robot interaction field trial conducted with the autonomous mobile robot *ACE* (*Autonomous City Explorer*) in a public place, where the *ACE* robot needs the support of human passers-by to find its way to a target location. Since the robot does not possess any prior map knowledge or GPS support, it has to acquire missing information through interaction with humans. The robot thus has to initiate communication by asking for the way, and retrieves information from passers-by showing the way by gestures (pointing) and marking goal positions on a still image on the touch screen of the robot. The aims of the field trial where threefold: (1) Investigating the aptitude of the navigation architecture, (2) Evaluating the intuitiveness of the interaction concept for the passers-by, (3) Assessing people's willingness to support the *ACE* robot in its task, i.e. assessing the social acceptability. The field trial demonstrates that the architecture enables successful autonomous path finding without any prior map knowledge just by route directions given by passers-by. An additional street survey and observational data moreover attests the intuitiveness of the interaction paradigm and the high acceptability of the *ACE* robot in the public place.**

***Keywords*-autonomous mobile robot; human-robot interaction; field trial; social acceptance;**

I. Introduction

Robots are gradually moving from structured industrial settings into our daily lives. However, current robotic systems are still largely tailored to their specific purpose and lack capabilities to adapt to their environment. A long-standing vision in the field of autonomous robotics has been to create systems that are capable of assisting humans in intelligent and versatile ways [1]. Impressive progress has been made in the field of unmanned outdoor navigation in unstructured terrains [2] and more recently in urban environments [3], [4]. There, autonomous robots have been taught to safely and reliably navigate in re-enacted traffic situations, avoiding collisions, staying on track, and respecting traffic rules.

However, all of these systems were provided global waypoints in the form of GPS coordinates, as well as topological information about the route in advance. In contrast, equally impressive achievements have been made in the field of indoor human-robot interaction (HRI). Prominent examples are tour guides for museums [5], [6] and shopping malls [7] that successfully relay useful precompiled information to humans.

But what happens if a robot is dependent on the support of humans to navigate in an unstructured outdoor environment? In this paper a human-robot interaction scenario is proposed, in which a robot is injected into a densely populated public place without any previous topological information and has to navigate to a target point provided only with the information it gets from passers-by and its proprioception. To assess the feasibility of this scenario, a field trial was conducted in a public place without instructing passers-by prior to their contact with the robot. The goal of this trial was to assess (1) the navigation architecture, (2) the interaction paradigm, and (3) the social acceptability of the robot.

The paper is structured as follows. Following a quick survey on related work in Sec. II, the navigation concept and implementation of the Autonomous City Explorer are presented in Sec. III. Sec. IV discusses – based on a description of the human-robot interaction scenario – the methodological approach and the instruments used in the field trial. Results and insights gained and recommendations for future work will be given in Sec. V.

II. Related Work

Systems assisting humans in a cognitive way will not always possess all the information required to fulfill their task. A central aspect of intelligent autonomous behavior is thus the ability to retrieve information through interaction. The great majority of HRI research is based on the assumption that the robot has complete knowledge of its environment. The few exceptions include human-augmented mapping [8] which enables a robot to create a map of its environment by exploration and asking a human to label the areas of interest. The robot Biron [9] integrates spoken dialog and visual localization to learn and label new places. All of these robots only use human-robot interaction as a means to attach human-understandable labels to their spatial belief.

Robots that ask humans for directions in order to extract information about their environments are all still operating in very structured indoor environments. A wheelchair robot [10] can be given coarse qualitative route descriptions. The office robot Jijo-2 [11] can learn the locations of offices and staff by moving around and asking humans for information. A robot asking for the way at a robotics conference is presented in [12]. A global inference approach [13] aims at enabling a robot to autonomously find a path within an office environment based on human directions. A miniature robot that can find its way in a model town by asking for directions is described in [14]. This robot is the only one employing natural language communication to retrieve its information.

Human-robot interaction scenarios for robots navigating in public spaces and interacting with naive users furthermore have to take into account several issues, like the ideal proximity between the human and the robot [15], [16], how to initiate the interaction with humans [17], [18], and how to communicate route directions [19]. However, little effort has been put on the constraint that the interaction between the user and the robotic system in public places is mainly short-termed, as passers-by normally will follow another

978-1-4244-4893-7/10/$25.00 © 2010 IEEE

action goal than interacting with a robot. Thus, for a sustainable human-robot interaction in public places intuitive input modalities play a major role.

III. The Experimental Scenario

A. The Autonomous City Explorer – ACE

In its current setup, the *ACE* robot comprises a differential drive mobile platform, developed by BlueBotics SA, two laser range finders, a speaker, a touch screen, an animated mouth, as well as a stereo vision system based on a multi-focal active camera head. Fig. 1 shows the robot and its components. The complete system measures 178 cm in height, including the camera head, and weighs approximately 160 kg. A more detailed description of the hardware architecture and the system design can be found in [20], [21].

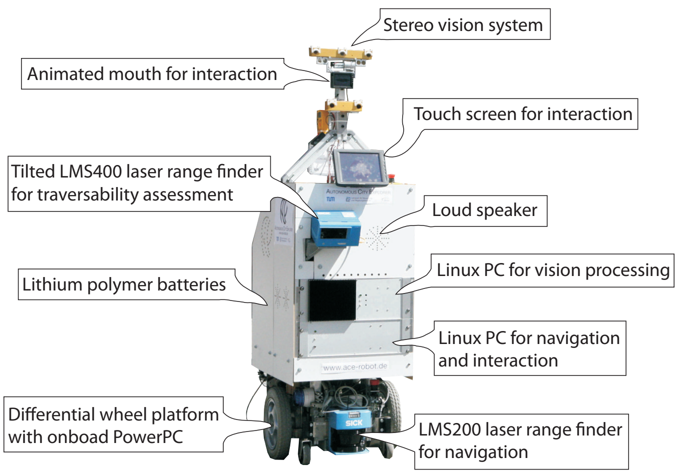

Figure 1. *ACE* with its main components.

The mobile platform with a laser range finder serves as the mobile base of the system. Using particle filter-based SLAM, the robot builds a $30 \times 30\,\mathrm{m}^2$ occupancy grid of its surroundings. In this grid it employs a dual path planning strategy using both a bounding box-based visibility graph planner for open spaces such as pedestrian areas and a Voronoi-based visibility graph planner for narrow passages such as sidewalks.

To avoid falling down the sidewalk or other negative obstacles, *ACE* uses a downward-looking laser range finder mounted at a 30° angle to assess the traversability of the terrain in front of it. The information from this laser is fused with the occupancy grid used for navigation. During interaction the robot stands still and the *Navigation* system is in idle state. Further information on the *Navigation* module of the robot can be found in [22].

The *Vision* system consists of a multi-focal actuated camera head. To find prospective interaction partners *ACE* detects humans using a combination of simple algorithms that yield robust and real-time capable. Given the common characteristics of pedestrians, such as skin color, motion, and upright posture, a top-down biased, task-relevant saliency map model is applied that predicts the positions of passers-by in an input image.

The actuated camera head turns towards the location of the most salient point, i.e. one of the detected pedestrians. The movement of the camera head expresses the robot's current interest in interacting with one of the pedestrians.

The *Vision* module also recognizes gestures during interaction with humans. The gesture recognition algorithm takes possible locations of humans as input, computes a disparity map containing depth information, this results in a three-dimensional representation of the scene. Starting with the detected skin parts, the algorithm segments the point cloud into smaller clusters and fits a 28 degrees of freedom kinematic human model into the scene, thus retrieving the body pose and the direction of the pointing gesture. For further information on the *Vision* module please refer to [23].

B. Interaction Paradigms

ACE interacts via speech and image output and gesture and touch screen input. The interaction is controlled by a finite state machine, that interfaces with the *Navigation* and the *Vision* modules, depicted in Fig. 2. The interaction is structured according to the four phases *Introduction*, *Giving Directions*, *Confirmation*, and *Conclusion*, usually applied in a similar situation in human-human communication. After greeting and explaining the task, the robot asks people to point in the first direction on the way to the designated goal location. The procedure of asking for a gesture as a first direction information ensures that both the human and the robot adopt a common reference system. As soon as both partners have a common reference system, further directions can be given unambiguously to the robot via buttons on the touch screen. During the interaction *ACE* builds a topological route graph from the presented route information. At the end of the interaction the robot thanks the passer-by and starts moving away along the route graph. As it makes progress, it associates the nodes in the topological route graph with the metric occupancy grid and updates the data. Further information on the *Interaction* module can be found in [24].

C. Environment and Task

The interaction context for the *ACE* robot is an outdoor scenario. In the specific scenario of the case study presented in this article the *ACE* robot has to reach Marienplatz (the central square of Munich), starting at the Odeonsplatz. The robot has no prior map knowledge or GPS sensors, and therefore has to ask passers-by in order to complete its task. The distance *ACE* had to cover was approximately 600 m in a crowded pedestrian area.

IV. The Field Trial

A field trail was conducted on September 2, 2008 in the city center of Munich, Germany, to investigate the aptitude of the navigation architecture, the intuitiveness of the interaction paradigms, and people's willingness to support the *ACE* robot in its task, i.e. the social acceptability. Following linguistic principles from human-human interaction as discussed in [24], a model interaction is sketched as follows: *ACE* asks for help "Please touch the screen in order to help me", a passer-by is ready to help *ACE* and touches the screen. *ACE* says "Hello, my name is *ACE*, I am looking for the way to Marienplatz. Please look into my eyes and point in the direction I should take now, as shown on the screen". After the interacting person shows the direction and *ACE* recognizes it, a still picture of the designated direction is taken and the person is asked to mark the correct goal position in it in order to verify the information. In a next step, *ACE* asks "How far is it to Marienplatz?" The estimated distance in meters is entered by the human on the touch screen. When *ACE* has gathered all this

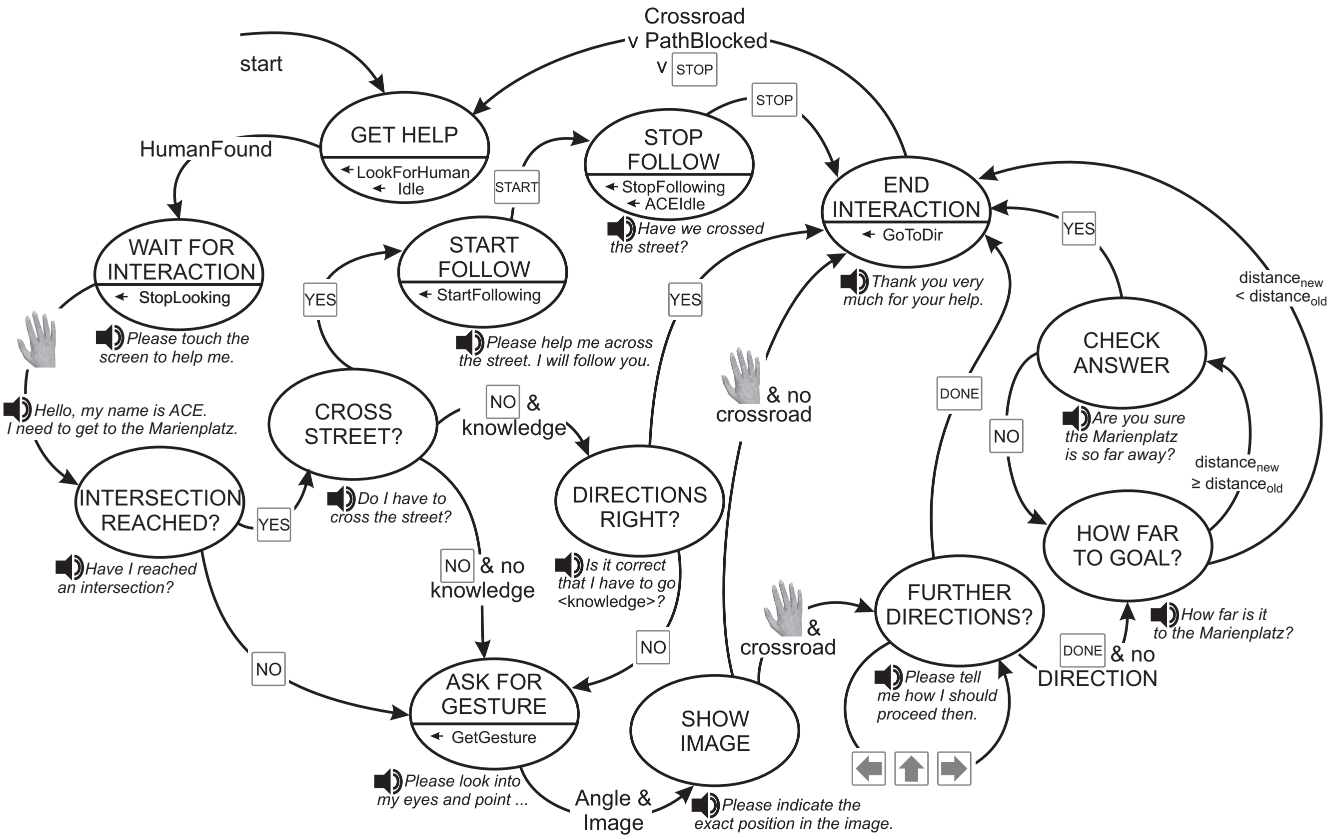

Figure 2. Finite state machine of the *Interaction* subsystem.

information it thanks the person and says: “Please step back I am going to go now”.

A. Study Setting

The field trial was conducted as a follow-up study to a previous wizard-of-oz experiment conducted in 2008 [25], where the robot was remote controlled. In the presented field trial the scenario was more complex and the robot exhibited more sophisticated abilities, such as autonomous navigation. The study was conducted in the style of a so called “breaching experiment”, which is a social psychological method to “detect some expectancies that lend commonplace scenes their familiar, life-as-usual character, and to relate these to stable social structures of everyday activities” [28]. The main difference of the methodological approach to standard HRI field trials is that participants are not previously informed in any way about the introduction of a robot into their everyday life. Since one goal of the experiment was to investigate the social acceptance of the robot by simulating its integration into human society, “breaching experiments” ensure a most natural and intuitive behavior of participants of the experiment. However, to ensure the safety of the passers-by during the whole user study, researchers (unidentifiable for the pasers-by) from the Technische Universität München were nearby to remotely control the locomotion of the *ACE* in case of emergency.

B. Instruments and Data Analysis

The study setting was based on a combination of quantitative measures – a street survey conducted with a standardized questionnaire – and qualitative data gathering and analysis – observational data interpreted by means of interaction analysis technique (explained in more detail in the following).

1) Observations: The aptitude of the navigation architecture was investigated during the run to Marienplatz by monitoring the status of the robot and its decisions and actions remotely on a laptop. Additional observation data was gained by an additional video camera mounted on the head of *ACE*, which recorded passers-by activities on the touch screen and their comments and discussions about the robot. Furthermore, one researcher accompanied the whole field trail to take observation notes on eye-catching or repetitive behavioral patterns.

2) Street Survey: To investigate the intuitiveness of the human-robot interaction paradigms and to assess the social acceptability of the *ACE* robot as an actor in the public place, a street survey questionnaire was developed consisting of seven questions on “Intuitiveness of Interaction”, five questions on “Sociability of Interaction”, nine questions on “(Social) Impact”, and two questions on “Approach Distance”. All questions had to be rated on a 5 point Likert scale from 1: absolutely disagree to 5: absolutely agree. The questionnaire served as self-reporting instrument giving insight on the *autonomous* perception of passers-by, i.e. their personal experience of the interaction with the robot.

The questionnaire based interview started with a short introduction based on the Thomas theorem [26] stating: “If people define situations as real, they are real in their consequences”. The introduction text sounded as follows: “You have just seen the *ACE* robot. It is used by the Technische Universität München as an errand-boy. With the help of passers-by it asks its way through the city and carries out a variety of catch and carry tasks for employees of the University. We would like to know, what you think about the *ACE* robot and ask you to answer some questions.” By assigning the robot a real-world task it was intended to support the illusion of *ACE* being a social actor and therefore stimulate participants to answer the questionnaires under the assumption that the robot is an active part of the social order.

The street survey was conducted with passers-by selected by chance by four researchers. This way of non-probabilistic sampling, where the researchers choose their participants by themselves is typical for street survey, as the target population can hardly be defined in advance. The rational choice for the researchers was only limited by one criterion: Only passers-by should be chosen who moved in a spatial context in which they could have noticed the robot. A disadvantage of this sampling method is that non-probabilistic sampling does not guarantee to address all relevant representatives of a target group, e.g. unfortunately in the presented field trial no passers-by younger than 20 years could be interviewed.

3) Interaction Analysis: Based on the recorded video data an interaction analysis following [27] was conducted. An interaction protocol was kept including the duration of the interaction as well as the conversations and gestures when people talked with *ACE* or with each other about *ACE*. The conversations were deeply analyzed by means of “fine structure analysis”, a method where the text is split up into small sequences that are chronologically analyzed in terms of latent meaning, intention, everyday meaning, distribution of roles, and options for the following sequence [27]. For this purpose, the text-sequences are decomposed into small units of meaning, while each unit of meaning should comprise at maximum one sentence. Afterwards every unit of meaning is analyzed in accordance to the following questions:

1) Paraphrase: What is the superficial information of the unit of meaning, what does it mean in every day life?
2) Intention: Which functions and intention could the producers of the text have had? What is the person going to achieve with the statement?
3) Latent Meaning: Which latent moments can be found in the unit of meaning?
4) Roles: Which distribution of roles is given through the statement?
5) Connection: Which options appear for the next unit of meaning?

After a separate interpretation of the units of meaning by five researchers (the method requires that at least two researchers do the interpretation, but the more researchers interpret the data the reliability increases), they are composed again. At the end emerges a consistent “structure of sense” containing the conditions, which have led to the production of the text and the significant meaning for the social actors in a social system, hidden “between the lines”. This interaction analysis served as self-reporting instrument characterizing the *heteronomous* perception of the interaction of passers-by with the robot, i.e. the interpretation of the interaction process by others.

Table I
EXAMPLE OF THE INTERPRETATION OF A UNIT OF MEANING

Paraphrase	Intention	Latent Meaning	Roles	Connection
Change position to observe a person	I want to find out what he is doing (curious)	I cannot see anything from here	father and child	no

Following an example text-sequence from a single person is presented already split into its units of meaning. (1) Let’s go over there to watch what he is doing, (2) he stops, (3) yes he stops watching,(4) he reacts on obstacles,(5) he drives around obstacles. An extract of the interpretation of the first unit can be found in Table I.

C. Results

The robot reached its goal after about two hours and interacted with 52 passers-by. The large number of people interacting arises from the fact that many of the interactions were started by curious passers-by. This also explains the relatively long duration of the experiment. Independently from the 52 passers-by interacting with *ACE*, randomly 52 passers-by were chosen to participate in the street survey (it was a coincidence that exactly the same number of passers-by were interviewed as showed *ACE* the way). From the interviewed participants 39 were male and 13 participants were female. The age ranged from 20 to 75 years with a mean age of 48.16 (SD: 16,78). Additionally participants were asked to state their educational level: 14 participants finished compulsory school, 10 graduated from vocational college, 6 finished middle school, 8 finished secondary school, and 8 had a university degree (3 participants stated other educational levels).

From the 52 interviewed passers-by 23 stated that they only noticed the *ACE* robot, 29 also interacted with the robot and non of the participants stated that they did not notice *ACE* (which is an indicator for the novelty effect and that the robot “breached” the everyday order of the passers-by).

1) Aptitude of the Navigation Architecture: The chosen architecture enabled the robot to successfully travel through the city of Munich, see also a previous experiment described in [21]) and interact with many people. An example of an interaction between *ACE* and a passer-by is shown in Figure 3, where a passer-by points into the first direction the robot has to go while *ACE* follows the gesture with the camera head to take an image and present it to the human.

Figure 3. *ACE* following the pointing gesture of a passer-by with the camera head.

Figure 4 depicts an occupancy grid generated by the robot during its run from proprioceptive laser-range finder data. To demonstrate the accuracy of the environment representation, a satellite image of the environment is overlaid. During interaction with humans the robot formed spatial knowledge and represented it as a route graph, an example of which is shown in Fig. 4.

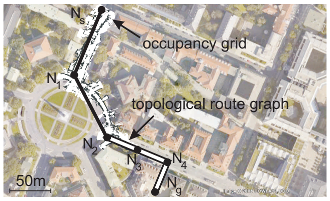

Figure 4. Example of a route graph retrieved from human instructions.

The fact that the robot reached its goal solely with the help of instructions of passers-by who were not previously instructed on how to interact, leads to the conclusion that the interactions were successful and fulfilled their purpose. Problems arose, where the human partners had too high expectations of the abilities of the robot. For example many users expected the robot to be able to understand speech at first and tried to answer through natural language until they realized that they had to use gestures and the touch screen to provide information. Furthermore, some humans encountered problems making a pointing gesture that was recognizable by the robot, as the robot could only recognize gestures where the arm was fully extended and failed on “sloppy” pointing gestures. The robot was sent in the wrong direction once, but this wrong information was corrected by the next interaction partner. Otherwise no conflicting information occurred.

To summarize most of the problems in human-robot interaction were related to unclear or missing statement of abilities and constraints in the communication process. Therefore a robot must clarify those points at the beginning of every new interaction. The next step is to make the communication more natural, in specific improve the gesture recognition and include a robust speech recognition system incorporating all of the introduced heuristic rules for interpreting deictics.

2) Intuitiveness of the Interaction Paradigms: The duration analysis based on the observational data revealed that the time until people give up interacting with the robot is rather short, the first impression and experience is decisive. If human and robot fail to establish a communication process within one minute, passers-by lose interest (“I don’t care any longer, let’s leave it alone . . . ”) or even verbally attack the robot (e.g. after a few trials a person says: “does not figure it out. . . ” and shows the robot with a gesture that it is ’a bit under the weather’). The mean duration of the interaction with *ACE* was 63 seconds (SD: 34.96); successfully completed interactions took 88 seconds in average (SD: 29.47), while an aborted interaction took 49 seconds in average (SD: 30.61). If the interaction works out immediately, people are even willing to stay and finish the interaction, but if it does not work as expected they do not hesitate to break up the interaction.

The street survey revealed the mean values regarding the questions on “Intuitiveness of Interaction” depicted in Table II, which indicate that passers-by had the feeling that they would need additional training to successfully interact with *ACE* and that they do not believe in their abilities to interact with *ACE* under time pressure. However, a difference could be identified for passers-by who actually interacted with *ACE*. A Kruskal-Wallis test was significant for question 4 $\chi^2(1, N = 49) = 9.16, p = (0.002)$ and question 2 $\chi^2(1, N = 52) = 7.33, p = 0.007)$ revealing that participants who interacted with the *ACE* robot rated these questions significantly higher than participants who only noticed the robot. Furthermore, a difference could be found for question 1, regarding the age of the participants. We split participants into two age groups: younger than 50 years and older than 50 years. The second age group (older than 50 years) rated question 1 significantly higher ($\chi^2(1, N = 52) = 4.12, p = (0.047)$), indicating that there is still an anxious attitude towards technology among the older population.

Table II
QUESTIONS ON INTUITIVENESS OF INTERACTION

No.	Question	N	Mean	SD
1	I am afraid of making errors if I am interacting with *ACE*.	52	4.23	1.37
2	I could work with *ACE* if I got a good training.	52	4.56	0.96
3	Under time pressure I could never be successful in dealing with *ACE*.	50	3.10	1.34
4	The interaction with *ACE* is easily understandable.	49	4.00	1.06
5	I could solve all occurring interaction problems with *ACE* by myself.	43	1.86	1.23
6	It will be easy to use *ACE*.	48	4.31	0.97
7	It will be easy for me to skilfully interact with *ACE*.	51	3.98	1.21

3) Social Acceptability of the ACE *robot:* The introduction of *ACE* into public space triggered a lot of conversation on the robot as well as different emotional reactions: curiosity, joy, enthusiasm, amazement, but also (not that often) fear and anger. People asked many questions – an indicator that the robot “breached” the everyday situation. To encounter *ACE* in a public place was an undefined and new situation for the passers-by triggering them to try to find out what other people thought about it and in which way other people explain and dealt with the situation. Nevertheless people showed different reactions to *ACE* (both positive and negative). *ACE* attracted nearly all passengers’ attention in some way, only few people showed no interest by demonstratively ignoring *ACE*, looking bored or busy, and similar ways of disinterest. This happened especially when *ACE* stood still and did not speak.

The observational data revealed five different types of passers-by, the first three being passive, the last two active:

Type 1: Ignores the robot and goes on without any reaction but showing disinterest (passive).

Type 2: Stares at *ACE* with eyes wide open, the way that young children look at people before they have learned that staring at someone is against social rules, and tries to see what is going on, but does not stop because of the robot (passive).

Type 3: Stops and watches what is going on with the robot from a safe distance and observes what other people are doing with it (passive).

Type 4: Talks to other people sometimes even to people they apparently did not know before to exchange opinions and definitions on *ACE* (active).

Type 5: Goes directly up to *ACE* and checks out by trial and error what the robot is for and how to handle it. For this type of people the robot presents some kind of riddle that has to be solved (active).

Of course people partly also mixed up the methods of gathering information and defining situation methods, e.g. asking somebody questions on *ACE* before directly interacting with *ACE*.

In the street survey two statements regarding the “Approach Distance” had to be rated by the participants: (1) I felt threatened because of the presence of *ACE*. (2) It is okay for me if *ACE* approaches me in the following distance .... For the second statement participants were asked to indicate their approach distance on the following social distance diagram (see figure 5). This social distance diagram depicts four approach distances: intimate, personal, social, and public which are defined by Hall in his book “The Hidden Dimension” [29].

1) Intimate space is the area immediately surrounding the individual’s body. This area is the most private one and it involves both, physical and emotional interactions.
2) Personal space is the area within which a person only allows selected friends or fellow workers with whom personal interaction is mandatory.
3) Social space is the area within which the individual expects to make purely social contacts on a temporary basis.
4) Public space is the area within which the individual does not expect to have direct contact with others.

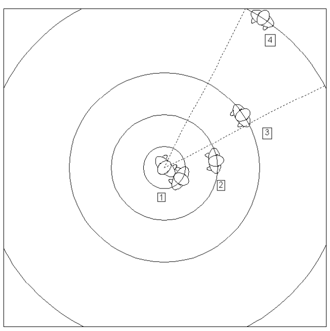

Figure 5. The Social Distance Diagram

The first statement was rated by the participants with 4.52 in average (SD: 1.26) which indicates that the presence of the *ACE* robot alienates passers-by (a further indicator for the “breaching” of the everyday order). On the other hand regarding the second statement 37 participants stated that they would allow *ACE* to enter their intimate space, 12 participants stated that they would allow it into their personal space, and only three participants chose the social space as an answer. This shows the willingness of people to interact with the *ACE* robot hazarding the consequence that they have to approach the robot closely to do so.

For the questions on “Sociality of Interaction” the street survey revealed the mean values shown in Table III. Comparing the results of questions 2 and 3 indicates that people did not experience the interaction as social as they desired it to be. However, significant differences could again be revealed for participants actually interacting with *ACE* and participants who only noticed *ACE*. All questions were rated significantly higher by participants with interaction experiences (Kruskal-Wallis Test: all $\chi^2(1, N \geq 41) \geq 5.68, p \leq (0.014)$). Moreover, questions 1 and 3 were significantly better rated by the second age group (participants older than 50 years), indicating that elderly people experienced a higher reciprocity in human-robot interaction (Kruskal-Wallis Test: all $\chi^2(1, N \geq 41) \geq 4.90, p \leq (0.027)$).

Table III
QUESTIONS ON SOCIALITY OF INTERACTION

No.	Question	N	Mean	SD
1	*ACE* reacted to my behavior.	48	3.31	1.72
2	The interaction with *ACE* should be a give and take.	51	4.02	1.36
3	The interaction with *ACE* was like a give and take.	41	3.15	1.68
4	*ACE* and me would compose a good team.	49	3.57	1.31
5	*ACE* needs my support to carry out its task.	48	3.83	1.62

The mean values of the answers to questions on the “Social Impact” are given in Table IV. Several interesting differences in the data regarding the independent variables, age, education, and if participants actually interacted with the *ACE* robot could be identified. The second age group (participants older than 50 years), rated the estimated effort it will make to interact with *ACE* in everyday life higher, than the first age group (participants younger than 50 years), see question number 2 (Kruskal-Wallis Test: $\chi^2(1, N = 49) = 6.19, p = (0.013)$). Participants who only noticed *ACE* rated question 7 significantly higher than participants who actually interacted with *ACE* (Kruskal-Wallis Test: $\chi^2(1, N = 49) = 10.78, p = (0.010)$). This result allows two interpretations: First skeptical passers-by did not interact with *ACE* and thus rated the statement lower; secondly participants who interacted with *ACE* where positively influenced by the interaction and thus rated this statement higher, which indicates that the interaction positively influenced people in their opinion about the robot. Subsequently, question number 8 was significantly higher rated by participants who interacted with *ACE* (Kruskal-Wallis Test: $\chi^2(1, N = 51) = 5.00, p = (0.025)$). An interesting finding was that participants with low education (only compulsory school) rated question 6 the highest (Kruskal-Wallis Test: $\chi^2(5, N = 50) = 10.17, p = (0.038)$), which affirms the assumption made by Arras and Cerqui [30] that a society without work is appreciated by low-skilled people who do not consider work as a joyful part of their life.

The interaction analysis revealed that passengers identified *ACE* as a robot. Many people used the term “robot” when recognizing *ACE*, which indicates that the embodiment of *ACE* fits to the mental model of what people think a robot has to look like (‘e.g. “Look there is a robot” or “a robot, a robot” or “roooobooot”). Moreover, not all passers-by believed that *ACE* was autonomous. The interaction analysis showed that it was not evident for all passers-by that autonomous robots exist which are able to move in public places without a human controlling them. Some passers-by tried to create evidence to verify their thesis that *ACE* must be controlled by a human. “There is someone sitting inside” or “Who is guiding it? It is that woman in the background, over there ...”.

Furthermore, passers-by tried to understand what the task of *ACE* was by adopting its perspective: “it is broadly going around us”, “it is reacting to obstacles”, reasoning that *ACE* looks out for people and thus that the safety of passers-by is granted. *ACE* also stimulated self-expression behavior. Passers-by who interacted with *ACE* demonstrated high self-confidence in the interaction (”Yes I can”-effect), e.g. a woman directly approached *ACE* while she was eating a snack and directly started to type on the touch screen without a moment of reasoning.

However, it could also be observed that women were generally shyer and were often interrupted during interaction by men with comments like, “Are you ready?” or “ Not finished yet?”.

The acceptance of *ACE* as social actor could above all be found in companion like behavior and comments of passers-by e.g., “Let’s have a look. . . oh yes. . . common. . . take of robot” or “Again, Again, common . . . its a path finder, a path finder, yes he wanted to know how to reach Odeonsplatz, isn’t that cool?” or “Now he is fed up with us . . . ”. Some passers-by even addressed the robot directly and not in the 3rd person: “Stop here, please”.

Table IV
QUESTIONS ON SOCIAL IMPACT

No.	Question	N	Mean	SD
1	*ACE* will be useful to me.	49	3.41	1.47
2	The effort to carry out tasks together with *ACE* will be immense.	49	3,06	1,41
3	I can imagine to carry out tasks faster with *ACE*.	51	4.02	1.22
4	I do not think it is necessary to introduce *ACE* to the everyday working life.	51	3.51	1.54
5	*ACE* could support me in everyday transactions.	50	4.18	1.10
6	*ACE* will ease burdensome tasks.	50	4.00	1.26
7	I would not like to work together with *ACE*	49	4.02	1.28
8	I think it is a good idea to “use” the *ACE* robot.	51	3.98	1.29

The interaction analysis revealed indicators for the willingness to accept *ACE* as well as the unwillingness. For instance one passer-by expressed contempt: “This is an autonomous robot that is going to crash the window pane”. Otherwise a lot of passers-by were happy to help *ACE* and went to interact with it on a voluntary basis. They interacted with it in a public place in a high frequented shopping street, where a lot of people watched the scenario, which is a strong indicator for acceptance. In general a positive social acceptance and impact of the *ACE* robot can be assumed based on the empirical data derived from the field trial.

V. CONCLUSION AND OUTLOOK

Are passers-by willing to support a robot, which is asking for the way in a public place? In this paper a field trial is described in which this question was investigated. The *ACE* robot (*Autonomous City Explorer*) was developed for an outdoor scenario in which the robot navigates to a designated goal location without any prior map knowledge or GPS sensors, only by asking passers-by for the way. The human-robot interaction was based on speech and gestures and touch screen input. In a field trail 52 passers-by interacted with the *ACE* robot to give directions, 52 passers-by were interviewed in a street survey, and more than 80 passers-by were video-observed while standing in front of the touch screen.

It could be shown that the navigation architecture enables successful autonomous path finding. Furthermore, the study revealed that the interaction paradigm chosen for this human-robot interaction scenario was intuitive above all for children and supports short term interaction in public places. Passers-by did not have the feeling that they needed additional training to handle the *ACE* robot. Passers-by who actually interacted with *ACE* even rated the statement “It will be easy for me to skilfully interact with *ACE*” on a scale from 1 = totally disagree to 5 = totally agree with mean: 4.39 (SD:0.96). Regarding the social acceptance and impact of the *ACE* robot it can be said that *ACE* attested a high acceptability rate in the survey data as well as in the observational data. Similarly, the fact that *ACE* reached its goal shows that people were willing to support *ACE*. The statement “It is a good idea to use the *ACE* robot was rated with a mean value of 3.98 (SD: 1.29).

Three main implications for human-robot interaction in public places can be derived from the field trial: (1) Supporting natural communication by structuring communication phases in accordance to human-human communication patterns, (2) Enabling short-term interaction: the first successful communication experiences must be received by the user during the first minute of interaction (3) Offering direct contact interaction as input modality can be used in pubic settings (e.g. touch screen) as passers-by accept that a robot enters their intimate space for interaction.

However, results presented in this paper are limited to the participants culture and the task of the *ACE* robot. They provide evidence on how robots might navigate in public places as integral part with the support of human passers-by. Further work is required to generalize our results and to extend our knowledge of intuitive short term interaction paradigms for human-robot interaction scenarios in public places. Future work will be necessary on the immediate mental models passers-by acquire regarding the capabilities and limitations of the robot. Research in this area will help for instance to find solutions how to initiate interaction with passers-by without raising the expectation of the robot understanding natural language. Moreover, the influence effect of user’s age needs to be investigated in more detail. To avoid laboratory-based studies, future field trials are planned in public contexts mainly frequented by children or elderly people (e.g. next to schools, kindergartens or living districts of elderly people).To gain more insights on interaction paradigms which even more decrease the cognitive load of inexperienced users laboratory-based user studies are planned comparing different input and output modalities while measuring the work load of the participants by means of heart rate variability analysis.

ACKNOWLEDGMENTS

The authors would like to thank all researchers supporting the *ACE* project (in alphabetical order): Klaas Klasing, Georgios Lidoris, Quirin Mühlbauer, Florian Rohrmüller, Stefan Sosnowski, and Tingting Xu. We would like to thank all researchers supporting the field experiment (in alphabetical order): Roland Buchner, Daniel Schwaiger, and Daniela Wurhofer. This work is supported in part within the European Commission as part of the Robot@CWE project, see also www.robot-at-cwe.eu, and within the DFG excellence initiative research cluster Cognition for Technical Systems CoTeSys, see also www.cotesys.org.

REFERENCES

[1] R. Brooks, “New approaches to robotics,” *Science*, vol. 253, pp. 1227–1232, 1991.

[2] S. Thrun and et al., “Stanley: The robot that won the DARPA grand challenge: Research articles,” *Journal of Robotic Systems*, vol. 23, pp. 661–692, 2006.

[3] M. Buehler, K. Lagnemma, and S. Singh, Eds., *Journal of Field Robotics: Special Issue on the 2007 DARPA Urban Challenge, Part I.* Wiley Periodicals, Inc., 2008.

[4] Y. Morales, E. Takeuchi, A. Carballo, W. Tokunaga, H. Kuniyoshi, A. Aburadani, A. Hirosawa, Y. Nagasaka, Y. Suzuki, and T. Tsubouchi, “1km autonomous robot navigation on outdoor pedestrian paths,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2008.

[5] W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, “Experiences with an interactive museum tour-guide robot,” *Artificial Intelligence*, vol. 114, pp. 3–55, 1999.

[6] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. Cremers, F. Dellaert, D. Fox, D. Haehnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz, “Probabilistic algorithms and the interactive museum tour-guide robot minerva,” *International Journal of Robotics Research*, vol. 19, pp. 972–999, 2000.

[7] H. Gross, H. Böhme, C. Schröter, S. Müller, A. König, C. Martin, M. Merten, and A. Bley, “Shopbot: Progress in developing an inter-active mobile shopping assistant for everyday use,” in *SMC*, (2008).

[8] E.A.Topp and H.I.Christensen, “Detecting structural ambiguities and transitions during a guided tour,” in *Proceedings of the International Conference on Robotics and Automation*, 2008.

[9] T. Spexard, S. Li, B. Wrede, J. Fritsch, G. Sagerer, O. Booij, Z. Zivkovic, B. Terwijn, and B. Kröse, “Biron, where are you? enabling a robot to learn new places in a real home environment by integrating spoken dialog and visual localization,” in *IROS*, 2006, pp. 943–940.

[10] R. Müller, T. Röfer, A. Lankenau, R. Musto, K. Stein, and A. Eisenkolb, “Coarse qualitative descriptions in robot navigation,” in *Spatial Cognition II*, 2000, pp. 265–276.

[11] H. Asoh, Y. Motomura, F. Asano, I. Hara, S. Hayamizu, N. Vlassis, and B. Kröse, “Jijo-2: An office robot that communicates and learns,” *Intelligent Systems*, vol. 19, no. 5, pp. 46–55, 2001.

[12] M. P. Michalowski, S. Sabanovic, C. F. DiSalvo, D. B. Font, L. M. Hiatt, N. Melchior, and R. Simmons, “Socially distributed perception: GRACE plays social tag at AAAI 2005,” *Autonomous Robots*, vol. 22, pp. 385–397, 2007.

[13] Y. Wei, E. Brunskill, T. Kollar, and N. Roy, “Where to go: Interpreting natural directions using global inference,” in *Proceedings of the IEEEInternational Conference on Robotics and Automation*, 2009, pp. 3761–3767.

[14] S. Lauria, G. Bugmann, T. Kyriacou, and E. Klein, “Instruction based learning: how to instruct a personal robot to find hal.” in *Proceedings of the 9th European Workshop on Learning Robots*, 2001.

[15] M. L. Walters, K. Dautenhahn, R. Boekhorst, K. L. Koay, D. S. Syrdal, and C. Nehaniv, “An empirical framework for human-robot proxemics,” in *AISB2009: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction*. Edinburgh, Scottland: SSAISB (ISBN - 190295680X), 8-9 April 2009, pp. 144–149.

[16] E. Pacchierotti, H. Christensen, and P. Jensfelt, “Evaluation of passing distance for social robots,” in *ROMAN 2006: Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication*, Sept. 2006, pp. 315–320.

[17] N. Bergstrom, T. Kanda, T. Miyashita, H. Ishiguro, and N. Hagita, “Modeling of natural human-robot encounters,” in *IROS 2008: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008.

[18] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, “How to approach humans?: strategies for social robots to initiate interaction,” in *HRI ’09: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. New York, NY, USA: ACM, 2009, pp. 109–116.

[19] L. S. Koulouri T, “A WoZ framework for exploring miscommunication in HRI,” in *AISB2009: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction*. Edinburgh, Scottland: SSAISB (ISBN - 190295680X), 8-9 April 2009, pp. 73–80.

[20] G. Lidoris, K. Klasing, A. Bauer, T. Xu, K. Kühnlenz, D. Wollherr, and M. Buss, “The autonomous city explorer project: Aims and system overview,” in *International Conference on Intelligent Robots and Systems*, 2007.

[21] A. Bauer, K. Klasing, G. Lidoris, Q. Mühlbauer, F. Rohrmüller, S. Sosnowski, T. Xu, K. Kühnlenz, D. Wollherr, and M. Buss, “The autonomous city explorer: Towards natural human-robot interaction in urban environments,” *International Journal of Social Robotics*, 2009.

[22] G. Lidoris, F. Rohrmüller, D. Wollherr, and M. Buss, “The autonomous city explorer (ACE) project - mobile robot navigation in highly populated urban environments,” in *Submitted to the IEEE International Conference on Robotics and Automation (ICRA)*, 2009.

[23] Q. Mühlbauer, S. Sosnowski, T. Xu, T. Zhang, K. Kühnlenz, and M. Buss, “Navigation through urban environments by visual perception and interaction,” in *International Conference on Robotics and Automation*, 2009.

[24] A. Bauer, D. Wollherr, and M. Buss, “Information retrieval system for human-robot communication - asking for directions,” in *International Conference on Robotics and Automation*, 2009, pp. 4150–4155.

[25] A. Weiss, R. Bernhaupt, M. Tscheligi, D. Wollherr, K. Kühnlenz, and M. Buss, “A methodological variation for acceptance evaluation of human-robot interaction in public places,” in *Proceedings of the International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2008, pp. 713–718.

[26] W. I. Thomas and D. S. T. Thomas, *The child in America; behavior problems and programs*. New York: A.A. Knopf, 1928.

[27] M. Lueger, *Grundlagen qualitativer Feldforschung: Methodologie, Organisierung, Materialanalyse*. Vienna: WUV-Univ.-Verl., 2004.

[28] H. Garfinkel, *Studies in ethnomethodology*. Englewood Cliffs, N.J.,: Prentice-Hall, 1967.

[29] E. T. Hall, *The hidden dimension*. New York: Garden City.

[30] K. Arras and D. Cerqui, “Do we want to share our lives and bodies with robots? A 2000 people survey,” Tech. Rep., 2005 2005, nr. 0605-001.