Studying Drink-Serving Service Robots in the Real World*

Rosalyn Melissa Langedijk\(^1\) and Cagatay Odabasi\(^2\) and Kerstin Fischer\(^1\) and Birgit Graf\(^2\)

Abstract—Field studies where robots are tested in real life settings bring different challenges for researchers, robotics scientists and users. In this paper, we address some of the challenges we encountered when testing two different drink-serving service robots in the wild. We collect challenges from three different experiments. Two experiments were conducted in elderly care facilities, while a third experiment took place in the lobby of a concert hall. We focus on the challenges that researchers face during the preparation phase and the on-set deployment phase when testing robots in the wild. We point to potential difficulties that may arise and present some practical solutions to the issues encountered. Our results suggest that lab studies do not sufficiently prepare the researcher for research ‘in the wild.’

I. INTRODUCTION

Testing service robots in the real world remains a significant challenge despite the growing technology.

Human-assistive technologies are in high demand for overcoming the challenges of an ageing society. The goal of these technologies is to be of practical use and make lives easier for, for instance elderly people and/or their caregivers. In this paper, we focus on a specific task, namely serving drinks. Since dehydration is a common problem in elderly care, which is also supported by our own ethnographic fieldwork, robots could possibly support caregivers in encouraging residents in elderly care facilities to drink more.

However, several challenges arise when testing robots in the real world. In the current paper, we report on the challenges we encountered when employing our robots in drinking interactions in the wild and develop suggestions on how to address them. The challenges concern a wide range of organizational, technical and practical issues, which we observed in three different studies (described further in section III).

When deploying and evaluating robots in real environments, three different phases can be distinguished. The first phase is the preparation phase, during which researchers and roboticists figure out what, where and how to test, and in which they discuss their planned tests with representatives of the testing facility. In the second phase, the focus is on technical feasibility and user experience. In this phase, roboticists supervise the tests in order to ensure the correct operation of the robot, sometimes even remote controlling individual functionalities. The third phase concerns robots that have reached a certain maturity level. This allows testing for a longer period of time, leaving the robot in the facility and letting the staff work with it alone. In this phase, the evaluation does not concern technical issues but rather the usefulness of the robot, for instance, improving work efficiency or the quality of work. The experiments covered in this paper deal with the first two phases of deployment.

II. PREVIOUS WORK

Previous work, e.g. [1]–[3], shows the importance of testing robots in the wild. For example, Jung & Hinds [4] argue that field studies are necessary to study human-robot interaction in complex social settings, especially how robots influence the social dynamics in real-life situations. Studies in the lab cannot replace field trials, and over-reliance on controlled lab studies with single individuals in controlled settings leads to an oversimplified view of human-robot interactions [4].

Thus, more recently, robot development has moved away from user-centered design approaches, which take one-on-one interactions as a starting point, to more contextualized, socially situated design approaches, such as stakeholder-centered design [5] or integrative social robotics [6], which requires the interdisciplinary design team to understand the design task to be to design culturally situated interactions, not robots [7]. The complexities and dynamics of real-life situations pose however considerable challenges to the researcher who addresses these issues, and only little field testing has been carried out to address such issues. However, those studies that were carried out in real-world scenarios confirm the complexity of the social environments that could not be anticipated from lab results (e.g. [8]) and the influence of robots on the social dynamics of groups of people. For example, Nomura et al. [9] found especially groups of children to constitute challenges for a robot in a shopping mall, and Mutlu & Forlizzi [10] show how a robot placed in a hospital is received differently in different wards depending on the ambient stress level.

Previous work also reports on numerous technical challenges for robots in the wild; for example, they have to deal with cluttered, dynamic environments (e.g. [11]) and much higher dynamics on robot robustness than in the lab. Blond [12], for instance, find hardware and software problems, such as ‘frozen tablets’, usability problems and issues regarding multiple use cases in a long term field trial of a robotic brain fitness instructor in a Danish elderly care facility.

In addition to these issues, there are numerous challenges that concern the investigation of robots in the wild, and
there is hardly any previous work on the potential practical problems that researchers may encounter. In this paper, we describe challenges from the robotics scientists perspective and from the perspective of the human-robot interaction researchers, as well as from the perspective of the end-users, like the elderly residents of care facilities and their caregivers.

III. THE THREE EXPERIMENTS

The qualitative observations made in this paper are based on three experiments at three different sites, which all have the task of serving beverages in common.

A. EXPERIMENT 1: Danish Elderly Care Facility

This experiment consists of ethnographic fieldwork and pilot tests for the deployment of a robot in a Danish elderly care facility. The participants are staff and residents at this care facility, where older adults, often suffering from dementia, live in small groups of 5-6 residents per unit. We carried out ethnographic observation, focus group interviews and co-design workshops [13], where we focused on the three use cases transportation of laundry, serving beverages, and guiding the users within the building. Furthermore, we tested different functionalities of the SMOOTH robot (see III-D), such as its navigation and dialog system, as well as the use case implementations.

B. EXPERIMENT 2: University Canteen and Concert Hall

This experiment was carried out in the university canteen and in the lobby of the concert hall. During the day, participants were mainly university staff and students while in the evening, participants were members of the public who waited for admission to concerts or other public events.

While people were enjoying their lunch or waiting to enter the concert hall, the robot was driving around offering water to people. In total, we collected data during three lunches and four afternoon or evening events, where we recorded about two hundred interactions with the robot. The robot greeted people, offered water, told jokes and facts about water intake and finally closed the interaction with ‘cheers’ or a ‘goodbye’ (cf. also [14]).

C. EXPERIMENT 3: German Elderly Care Facility

The third experiment was carried out at a German elderly care facility, where the robot was operated in a common room in order to serve water, apple and orange juice. The field study lasted one week. The participants were mainly elderly residents, but we also carried out interviews with the staff. The common room includes tables, chairs, and a kitchen so that people can come together, eat and enjoy their drinks.

During the day, there are "rush hours" for the robot in the common room. They occur especially after breakfast and lunch. Additionally, there may be some social activities where the staff come together with the residents for special events. The robot is especially suitable during this time to attract people’s attention and to serve some drinks.

The three experiments were carried out in real-world environments in which the robots could be deployed in the future. EXPERIMENT 1 is a field study in which we observed how daily tasks are executed at a care facility and which provided us with initial insights into the workflow in elderly care facilities. EXPERIMENT 2 was a human-robot interaction experiment in the wild, where we elicited spontaneous interactions with a broad range of members of the public. Based on the results from these two experiments, we conducted EXPERIMENT 3 at an elderly care facility and thus involving members of the target audience. Together, the three experiments inform us on how to conduct human-robot interaction experiments in the wild and specifically in elderly care facilities. They also lead to similar challenges for the researchers, which we report on below.

D. Robots and Testing Procedure

For EXPERIMENTS 1 and 2, we used the SMOOTH robot [13], which is a large service robot developed to take over several tasks in elderly care facilities, including the transportation of laundry and guiding of residents, as well as serving water. The robot’s head includes a microphone, speakers, cameras and two touch screens, one in the front and one in the back. The front touchscreen displays a pair of simulated eyes. The robot is furthermore equipped with autonomous navigation and dialog capabilities; however, in the large field study in connection with EXPERIMENT 2, to ensure participants’ safety and to be able to adapt the dialog to the circumstances arising, the robot was controlled by two wizards in the field trial, one for navigation and movement, the other for the dialog. The wizards in EXPERIMENT 2 operated the robot from a balcony above the university lobby and were thus hidden from view for the concert attendees and visitors of the canteen. The drinks were stored on a tray on the back of the robot (see Figure 1).

For EXPERIMENT 3, we used the Robotic Service Assistant [15] developed by Fraunhofer IPA. The drink serving concept was investigated first in WiMi-Care project with Care-O-bot 3, which is a general-purpose service robot [16].
Since the scenario turned out to be promising and the conducted user tests showed positive results, the Robotic Service Assistant was built in the scope of the SeRoDi project. A complete redesign was necessary in order to provide a specialized, thus closer to a product, platform for the drink serving application. The robot consists of an omnidirectional mobile base, drink storage, serving mechanism, and a tablet for human-robot interaction. In Figure 2, a person is selecting a drink on the GUI of the robot. This image is from a two weeks test conducted within the SeRoDi project in another German elderly care home. The test results were used to improve the capabilities of the robot, such as the way a user is approached [17]. In the scenario depicted in the image, the robot is offering four different beverages. After the user has selected one of them, the robot starts handing over the respective beverage in of a cup. Another modification compared to the first tests was the integration of "eyes" on the tablet that are displayed while the robot is not serving to make it more "friendly." The robot fuses information from three lidars and one RGBD camera to ensure safety during the navigation. Other than for safety, the RGBD cameras provide data streams to the robot for several purposes, including people detection. Although the robot can autonomously navigate in public environments and approach a person, during EXPERIMENT 3, a wizard was used to operate the robot and a second wizard managed the dialogue. In this case, however, due to the structure of the application environment, the wizards could not be hidden from the participants’ view.

E. Robot Dialog

In all three experiments, the dialogues were scripted and played according to a dialog model, with some flexibility for the wizard. In particular, a set of functionally equivalent utterances was defined from which the wizard could choose in order to vary the robot output so that overhearers would not witness the same dialogues over and over again. There are:

- different greetings (e.g. hi, hello, sorry to bother you, but...),
- different robot utterances to offer some water (e.g. How about some water?, Would you like something to drink?, You are probably thirsty - please take a drink etc.),
- different persuasive utterances we were testing (e.g. Research shows that it is important to drink enough water during the day or Most women do actually take something to drink),
- different humorous utterances (e.g. What did the ice cube say to the water? I was water before it was cool),
- a request to take the water,
- a toasting utterance (cheers), and
- different closings (e.g. Enjoy your drink, It was nice meeting you, Have a lovely day).

In EXPERIMENT 1, the dialog was prepared in Danish, for EXPERIMENT 2 it was held both in English (during the day) and in Danish (in the evening), and in EXPERIMENT 3, the dialog was carried out in German. In all cases, the dialog models consisted of comparable utterances and sequences.

F. Data Analysis

The data analysis is purely qualitative and mainly based on ethnographic observation and interviews. The observations were generally carried out on-site, but in EXPERIMENT 2, they are also based on analyses of the video recordings.

IV. CHALLENGES IN THE PREPARATION PHASE

During the preparation of the experiment, the researchers define their research questions and figure out what to investigate and how and where to test them. Furthermore, they need to develop detailed time plans for the testing period, and organize the ethical approval for the experiments. The following sections give some insights into the specific challenges and requirements that we faced during this phase.

A. Interaction with Testing Facility

First of all, extensive discussion with the management and staff at the experimental sites, in our case, two elderly care facilities, is required. This is essential to understand the limits of the environment, the work flow into which the robot is to be inserted, the circumstances of the deployment etc. Informing the staff at the facilities is also an important aspect and crucial for the success of the experiment. Only if they know what this experiment is about, what they can expect from the robot and what its limitations are, they will support the experiment as needed. Specifically, they should be aware of how robots of this kind will be used once they reach a product level and that the intention is not to replace them but to support them with their work. This can be challenging if the personnel is changing, or if not everyone is supporting the deployment of the robot in the facility (cf. also [18]).

In the preparation of EXPERIMENT 3, we faced some challenges due to the geographical distance between researchers and care facility. In spite of quite extensive discourse with the facility management prior to the on-site experiments, not all staff and residents were involved in the preparation, which

Fig. 2. The Robotic Service Assistant is serving drinks in a care home. Details of the application can be seen in the video (from minute 2:55): https://youtu.be/dQ5p0h_p4M
may have led to some issues of the staff not giving full consent, which we describe in section Consent Forms below.

B. Consent Forms

Data protection legislation requires that participants in experiments provide informed consent. During our field trials, we faced different kinds of challenges in the three different scenarios. One challenge concerns the problem that collecting consent forms can be time-consuming and needs to be started well in advance if there are participants with cognitive challenges so that relatives need to be contacted in order to get consent. In EXPERIMENT 1, we faced further difficulties because the elderly are mostly people suffering from dementia, and since residents only live in this facility for approximately ten months on average, we had to collect new consent forms from new residents and their legal guardians repeatedly.

Much manpower is also necessary if consent needs to be gathered from many participants simultaneously, like in EXPERIMENT 2. One reason is that people may not understand the purpose and use of consent forms. People found it hard to comprehend why a researcher should need such a consent, because they are unfamiliar with the processes and requirements common in research. Consequently, when testing in the wild and collecting consent forms at the same time, we often had to explain to people what the forms mean and why they are important, which takes a lot of time.

In EXPERIMENT 2, where the participants were members of the general public, we also encountered people who did not want to participate and answered all questions in the consent form with ‘no’, including the question “I understand that my participation is voluntary and that I can withdraw my consent at any moment in time.” This means that researchers actually also had to briefly check every consent form people filled out. Since there were often many participants around simultaneously, finding everyone who interacted with the robot or was captured by one of the cameras, providing them with consent forms and guiding them through the process was very challenging.

One possible solution to the challenges of collecting consent forms is to ensure that there are enough researchers to collect consent forms when testing in the wild. Furthermore, it turned out to be helpful to have a short description of the purpose of the project at hand, whereas the consent forms themselves may be brief. For the first two experiments we used quite elaborate consent forms that had been created by our university’s lawyers, and for the last experiment, we translated these forms and adapted the project description. However, we experienced that most people preferred not to read very much, while additional information was useful to satisfy some more curious participants who tried to engage the researchers in discussions about robots in society in general. Furthermore, it turned out necessary to prepare for the necessity to explain terms like ‘research purposes’ or ‘scientific articles’.

In EXPERIMENT 3, the consent forms were sent to the elderly care facility in advance together with some additional information. The consent forms were passed from the staff to the residents and/or their relatives. Many residents signed the consent form and were willing to participate and give feedback; however, the staff was more reluctant to give a full consent. One possible reason for this could be that the staff didn’t know well enough what to expect, see section about Interaction with Testing Facility. This uncertainty can be addressed by personal meetings conducted prior to the tests where researchers explain to the staff directly what is going to happen.

C. Participants

Another, related, challenge concerns who the participants actually are in the respective experiments. For instance, in EXPERIMENT 3, consent was collected from the residents and staff on the floors where the experiments took place; however, it turned out that residents from other floors were also visiting, and there were relatives and other personnel who happened to be present during the days when the experiments were carried out, which could not have been anticipated. Thus, precautions have to be taken for unforeseen participants.

Similarly, challenges occurred if there were residents or other participants who did not want to be included in the experiments, but who happened to share the space with people who had signed up as participants. Since the elderly care facilities are actually their homes, one cannot simply close off an area only for participants who have signed the consent forms. Similarly, in EXPERIMENT 2, often many people gathered around the robot and interacted with it, of whom some may have signed consent forms whereas one may have refused to do so, which complicates the data analysis.

D. Technical Aspects

Concerning the technical aspects, from our point of view, it is essential that the robot has reached a certain maturity and dependability level before going for the on-site tests. In order to prepare for the on-site tests, extensive in-house testing and optimization of the robot is required. To do so, a realistic environment should be constructed in the laboratory; then, the robot should operate there without any problem. Furthermore, selecting the specific operation environment for testing in the facilities is required. In some cases, it might be necessary to adapt the environment to ensure that all the participants can be reached. Narrow passages or grouped chairs are typical challenges for which a solution needs to be found beforehand or during the tests. One of the biggest technical challenges that may arise in the preparatory phase is network usage because a functioning network connection is vital for operating the robot. Some robots may need internet access for logging, teleoperation, and cloud computing. If this infrastructure cannot be provided by the facility, the robotics scientists have to bring it with them and install it for the duration of the tests. Moreover, the robot has to come with sufficient error handling capabilities to ensure that it can cope with any unexpected changes in the environment.
V. CHALLENGES ON-SET

In the deployment phase, the researchers are moving their robot to the setting where the experiments are going to be carried out. Here, the first step is to install the robot in the environment, e.g., create a map and define target positions where users should be approached.

A. Introducing the Robot to the Participants

After the successful installation, pre-testing, and optimization of the test scenario, the robot should be introduced to the participants. This introduction should cover the two relevant target groups: staff and residents.

The care staff is an essential part of the tests. The robot is for helping them and for making their work easier. That is why they need to know the product idea, the full workflow, and how to use the robot. For example, the Robotic Service Assistant used in EXPERIMENT 3 has a GUI based on a smart phone for the staff to activate the scenario, call the robot, visualize its status (including level of drinks still available). In the scope of the experiments covered in this paper, the focus was on the resident. Therefore, the GUI was not used by the staff but by the researchers.

The residents interact with the Robotic Service Assistant through a touch screen attached to the frontal part of the robot. It turned out that some residents initially hesitated to interact with the robot since they did not know how to use the tablet. However, the experiments show that they get used to it and hence get comfortable over time with the robot as they learn more about it.

B. Voice

One major challenge, which we had not anticipated, concerns the voice of the robot. In the preparation for the dialog designs, we had used text-to-speech systems that we had judged as prosodically very advanced, yielding very naturally sounding synthetic speech. However, our older participants in EXPERIMENTS 1 and 3 had considerable problems understanding the robot’s speech, probably because of some hearing difficulties due to old age. In many cases, the researchers had to repeat every sentence the robot produced for the residents to understand. Furthermore, the robot’s voice in EXPERIMENT 2 was appropriate for the interactions with members of the public when the space was relatively empty. However, when more people gathered to wait for admission to the concerts and other events, the robot’s voice was not loud enough. One possible solution is thus to implement the ability to adapt the loudness of the robot’s voice depending on the ambient noise. Moreover, in EXPERIMENT 1, we found that when the robot was guiding a resident, it was hard to understand because the loudspeakers were facing forward while the elderly person is walking behind the robot. Thus, also the robot design needs to support the needs of the older participants.

C. Safety vs. Performance

Before testing the robots on-site, a risk assessment needs to be conducted and a risk mitigation concept needs to be designed. One essential safety element is the use of safety laser scanners to stop the robot when people get too close. According to applicable safety standards such as ISO 13482, the robots have to be equipped furthermore with red emergency stop buttons, which allow the operator to halt the operations of the robot directly or remotely. The SMOOTH robot and Robotic Service Assistant both have two of them, one on each side. In addition, a wireless emergency stop was provided to the person supervising the tests in EXPERIMENT 3. In EXPERIMENT 2, where participants were members of the general public, one researcher had to interfere and stop a group of kids from pressing this button. The wireless emergency device used in EXPERIMENT 3 showed some challenges as well. If the connection is interrupted, the robot’s operations will be terminated. Therefore, the remote control must have a robust connection to the robot. Finally, in order to be able to stop in time before a collision occurs, the maximum velocity of the robot should be limited, and it has to keep its distance to the people. This could be seen as a performance problem from the user’s point of view.

D. Refill Process/Serving Drinks

Several challenges occurred concerning practical issues of filling up the robot with sufficient numbers of drinks. First, the operators have to know when the robot needs to be refilled (a challenge in EXPERIMENT 3). In that case, the robot would either return to a previously defined refill position where empty cups are replaced by full ones (Robotic Service Assistant), or a researcher would exchange the cups during use (SMOOTH robot). With the Robotic Service Assistant, the beverages are safely stored inside its body, which has significant advantages regarding hygiene. On the other side, it is not possible to directly see when the robot needs to be refilled. Instead, a smart phone interface is used to provide this information to the staff. It also enables to easily update the status after refilling and to start or stop the operation of the robot.

In contrast, the beverages are all in full view for everyone on the SMOOTH robot, and the robot can be easily refilled by researchers just placing more glasses onto the robot’s tray. However, people can also, and were observed to do so, place their empty glasses on the tray as well, so that it is not always possible to say which glasses have been used already, and in general, this way of serving beverages has hygienic disadvantages.

When serving larger groups, the robot may not have enough beverages for all participants, in these instances, people may be disappointed that they are not being served, and it may lead to the perception of inequality if someone does not receive a beverage whereas someone else did. In this case, it is perhaps desirable if people can see the reason for the fact that they are not being served, or if the robot can let them know that it needs to be refilled first and will come back later. In EXPERIMENT 2, several participants tried to verbally inform the robot when its tray was empty.

A big issue in elderly care facilities is also the personalization of the beverages handed out. In EXPERIMENT 1,
we found that in the small units of Danish elderly care facilities, almost every resident receives special drinks in special containers; in EXPERIMENT 3, more general drink serving was possible, but it was still necessary to ensure, for instance, that a person with diabetes does not get the juice with sugar. In previous experiments, for this purpose and for documentation of the drinks given out, face detection was used to identify individual residents.

E. Ethical Challenges

Another challenge is that in two of the three experiments reported on the robot was remotely controlled for security reasons. In EXPERIMENT 3, the robot operator was in plain view of the participants, which led one participant to ask how he was supposed to trust the robot if the operator does not. In contrast, in EXPERIMENT 2, the wizards were hidden from the participants, and due to the kind of situation studied, a debriefing of the participants was not feasible. This can lead to an inadequate understanding of robots’ real capabilities and the state of the art in the general public (cf. [19]). The field researcher has to find a balance between these two challenges.

VI. DISCUSSION

In this paper, we presented several challenges we have faced when going into the wild with our robots. We divided the challenges according to when in the study they occur: preparation, testing with technical staff present and long-term testing without technical staff. The experience covered in this paper refers to the first two phases. Some of the challenges described may be prevented by site visits, ethnographic observation, a participatory design process and by getting a thorough understanding of the field site and the participants.

In addition, however, many challenges emerged that are due to the complexity and dynamics of real-life social settings, as Jung & Hinds [4] and Chang & Sabanovic [8] suggest. Studies in the lab cannot prepare the researcher for these challenges, some of which concern very practical issues, but others also reveal that currently our models of human-robot interaction are ill-equipped to deal with unforeseen, multiple participants, group dynamics (such as who gets what beverages) or ethical issues.

Our review of the previous work (see II, e.g. [2]–[4], [8], [10]) suggests the importance of testing in real environments instead of only testing in labs, however, mostly with a focus on the direct interactions. In comparison, we focus here on the challenges that researchers and robotic scientists face when going into the wild; we present our findings on aspects that need to be considered when designing and executing such studies.

ACKNOWLEDGMENTS

We are greatly indebted to the residents and staff at the Ølby care home in Køge and the Luise-Schleppe-Haus und Schloss care home of Stifting Evangelische Altenheimat in Stuttgart. We furthermore wish to thank Lotte Damsgaard Nissen, Oskar Palinko, Eduardo Ruiz Ramírez, Selina Sara Eisenberger, Matous Lelinek and Simon Baumgarten, who helped during the experiments.

REFERENCES