Int J of Soc Robotics (2015) 7:203–225
DOI 10.1007/s12369-014-0257-8

User-Centered Design and Evaluation of a Mobile Shopping Robot

Nicola Doering · Sandra Poeschl · Horst-Michael Gross · Andreas Bley · Christian Martin · Hans-Joachim Boehme

Accepted: 21 September 2014 / Published online: 28 October 2014

Abstract This paper describes the user-centered design and evaluation process of a humanoid mobile shopping robot named TOOMAS that assists customers of home improvement stores. Three separate empirical field studies addressing the robot's usability (according to ISO 9241-11) and acceptability (intention to use) are presented involving $N = 343$ test persons altogether. The first formative evaluation study ($N = 210$) addresses the usability of the robot's article search system. It is demonstrated how several usability problems could be identified and eliminated, leading to significantly more successful article searches. The second formative evaluation study ($N = 39$) addresses the robot's adaptation to its specific task and role in the home improvement store. Embodiment, mobility, voice output, and social behavior were analyzed and adapted to user requirements. The third summative evaluation study ($N = 94$) experimentally tested robot-assisted shopping against conventional shopping regarding core usability criteria (effectiveness, efficiency, and satisfaction). It reveals that robot-assisted shopping was as effective and satisfactory as conventional shopping. Still, at the current state of technology shopping with robot assistance was slightly slower and therefore less efficient. For all three studies the effects of users' gender, age, educational background and computer skills levels on robot usability and acceptability are presented and discussed.

Keywords Service robots · Shopping assistants · User-centered design · Usability · Acceptability

N. Doering · S. Poeschl (✉)
Institute of Media and Communication Science,
Media Psychology and Media Design Group,
Technische Universität Ilmenau, P.O.B. 10 05 65,
98684 Ilmenau, Germany
e-mail: sandra.poeschl@tu-ilmenau.de

N. Doering
e-mail: nicola.doering@tu-ilmenau.de

H.-M. Gross
Neuroinformatics and Cognitive Robotics Lab,
Department of Computer Science and Automation,
Technische Universität Ilmenau,
P.O.B. 10 05 65, 98684 Ilmenau, Germany

A. Bley · C. Martin
MetraLabs Neue Technologien und Systeme GmbH,
Am Vogelherd 22, 98693 Ilmenau, Germany

H.-J. Boehme
Faculty of Information Technology/Mathematics,
University of Applied Sciences Dresden,
P.O.B. 12 07 01, 01008 Dresden, Germany

1 Introduction

Service robots have seen a continuous increase in the last years. According to the International Federation of Robotics, more than 16,000 service robots for professional use were sold in 2012. Also, about 3 million service robots for personal or domestic use were sold in the same year [22]. Whereas professional service robots are used by experts, for example in military and agricultural application fields, personal service robots provide services for laypersons and are employed as guides in museums and exibitions or as household and shopping assistants [21,58].

During the last years, shopping robots in particular have been on the upturn (see Sect. 3). These are embodied computer-supported shopping assistants with the ability for autonomous movement within stores or malls [4,16]. The main goal of all robotic shopping aids is to make shopping in a supermarket or mall easier and more convenient for the broader public (and especially for elderly and disabled people). Typical functions of mobile shopping robots are managing the shopping list, providing product information, guiding through the shop and finding products, transporting the selected goods, and even serving as a social shopping companion. Their development, as well as the development of personal service robots in general, has to face a variety of technological challenges in the areas of navigation, integration, and human-robot interaction [17] (for typical tasks of shopping robots in mall-like infrastructures see also [62]). Moreover, shopping robot users as laypeople can only be trained to a limited extent [9]. Therefore, consideration of user factors throughout the whole design and development process is essential when developing shopping robots. Apart from the shoppers as end users, requirements of all target groups, including store owners and managers as well as staff, have to be considered.

This paper describes the user-centered design process of the humanoid mobile shopping robot TOOMAS based on three separate empirical evaluation studies:

The paper starts with a theoretical outline of robot usability and acceptability (Sect. 2) followed by a presentation of related work on shopping robots and their usability and acceptability (Sect. 3) as well as a presentation of the shopping robot TOOMAS and its features (Sect. 4). Then, the first formative evaluation study addressing TOOMAS ' article search system (developmental stage 1, Sect. 5), the second formative evaluation study addressing the robot's adaptation to its task and role (developmental stage 2, Sect. 6), and the third summative evaluation study comparing robot-assisted shopping and conventional shopping (developmental stage 3, Sect. 7) are presented. Results are discussed in Sect. 8 and conclusions are drawn in Sect. 9.

2 Usability and Acceptability in Robot Development

Human-robot interaction should be as intuitive as possible and therefore should acknowledge user needs of different target groups. Until recently, development of service robots appeared to be highly technology or need driven [56] in the sense that a robot's role was mostly that of a task completer. Of late, the perspective on a robot's role has broadended to being

> an Interaction Peer that can lead the interaction and resolve ambiguity in situations of naivety. [31, p. 139]

Therefore, a user-centered perspective is relevant in two respects. Firstly, the actual use situation: human-robot interaction should lead to task fulfillment in a user-friendly way (usability). These are essential prerequisites for the second aspect, addressing a long term perspective: service robots should be accepted by users (acceptability; [37]). We argue that usability is a key issue in user-centered optimization of personal service robots and shopping robots specifically, as those are usually used by laypeople without a dedicated training in handling robots. Regular customers of a specific shop should be motivated to use the robot repeatedly.

Although robot developers acknowledge the importance of testing prototypes for usability, most studies differ on which usability framework is used [66,72]. The reason for this might be that, as Yanco and colleagues state,

> current human-robot interfaces differ widely depending on platforms and sensors, and existing guidelines are not adequate to support heuristic evaluations [72, p. 120].

However, the frameworks developed for usability testing of graphical user interfaces can be transferred and adapted to evaluating human-robot interaction, as long as they take the complex, dynamic, and autonomous nature of robots into account [72]. The International Standards Organization norm ISO 9241-11 defines usability as follows:

> Usability is measured by the extent to which the intended goals of use of the overall system are achieved (effectiveness); the resources that have to be expended to achieve the intended goals (efficiency); and the extent to which the user finds the overall system acceptable (satisfaction) (ISO 9241-11, [51]).

Effectiveness, efficiency, and user satisfaction can be regarded as quality factors of usability. These dimensions have to be decomposed into usability measures [51], which can be divided into firstly, subjective usability evaluation methods like questionnaires and interviews. Those are directly based on the users' judgment, without requiring users to interact with a technological system. Secondly, objective methods – or observational methods like direct observation or video recording – involve real people using working systems and measure user performance in fulfilling a predefined set of tasks. Those objective measurements are independent of the users' perception [20,51]. By integrating subjective and objective measurements of usability, a comprehensive assessment is possible, as these may lead to different information and optimizations, for example not only improving objective performance, but also generating design advice to improve the user experience [20].

However, those usability criteria might in some cases be too narrow to gain detailed data, especially for optimizing a personal shopping robot and adapt it to its role, its task, and a specific environment. Lohse [40] developed criteria explicitly for designing personal service robots, which can be differentiated into the following dimensions:

- Adequate *embodiment* or appearance of the robot is of relevance [38]. As humans try to transfer human characteristics and emotions to machines [57], service robots showing human-like features may be experienced as more intutitively usable for laypersons [40]. Mori states however, that minimal discrepancies from real human characteristics can be disturbing, so the robot should not look overly human-like [46].
- *Movement and Mobility* Movement and speed of movement influence robot acceptability [24,40]. Abrupt changes of direction can be perceived as aggressive [6], and speed of movement should be adapted to user requirements [40]. At the same time, navigating in a home improvement store is a challenging task for mobile robots: during the guidance tour, they have to navigate around obstacles and stay in contact with users.
- *Voice Output* Information has to be presented in a way that is adequate for the interaction goal and the application environment. The physical environment of a shopping robot is challenging, as a home improvement store is a rather loud environment that leads to interferences: Voice output has to be easily understandable, optimizing intonation, loudness, and adequate speed.
- *Social Behavior* Bartneck and Forlizzi state that social service robots interact directly with humans [1]. Therefore, they should behave as human-like as possible and act according to social norms. Approaching customers should not be intrusive and take into account personal space [52,65].

Further, usability can be considered as a prerequisite for robot acceptability for personal service robots. Given the growing development and public interest in robotics,

> it is critical to understand the factors that may increase acceptance and adoption [2, p. 4].

Some researchers claim that a comprehensive theoretical model on robot acceptability is missing (for example [2]). However, technology acceptability has been widely researched up to now, and a number of technology acceptance models (for an overview see [70]) exist. They define *intention to use* and *using a system* as main dependent variables to characterize individual technology acceptability. Beer et al. [2] argue that general technology acceptance models are able to provide guidance for understanding the variables that influence robot acceptability.

In our project we evaluated a specific kind of personal service robot, namely a shopping robot in a home improvement store. We used usability criteria according to ISO 9241-11 for evaluating the article search system (Sect. 5) and for comparing shopping with robot assistance with conventional shopping (Sect. 7). We used Lohse's criteria in order to adapt the robot to user requirements considering its task and role (Sect. 6). For the usability tests, subjects were invited by interviewers to use the robot, as some criteria (like satisfaction) could only be measured after system use. Effects on future self-initiated robot use could not be measured due to time constraints. Instead, *intention to use the robot in the future* was inquired and regarded as an indicator of robot acceptability in the three studies. We limited our approach to user acceptability, as the project's main focus was the development of a user-friendly robot. Therefore, we will not discuss other acceptability factors like safety or legal regulations in this paper.

3 Related Work

During the last years, several teams of roboticists have presented new shopping robot prototypes, representing worldwide leading edge advancements in the field. The following account of shopping robots focuses on comprehensive presentations, which were published in scientific articles. Media reports were not included. As this paper focuses on a user-centered development perspective, related work will be presented from a human-robot interaction perspective. Issues of usability testing and robot acceptability will be discussed. Shopping robot prototypes can be grouped in *three broad categories*: robotic shopping carts, non-humanoid shopping robots, and humanoid shopping robots.

3.1 Robotic Shopping Carts

Nishimura et al. [50] implemented robotic hard- and software into a conventional shopping cart. The shopping cart follows customers autonomously and transports the goods. A camera and distance calculation enables the robotic shopping cart to follow users at an adequate distance. The robot can also be controlled by operators, no user study has been conducted.

Kohtsuka et al. [33] followed a similar approach: They equipped a conventional shopping cart with a laser range sensor to measure distance from and the position of its user and an evasion system to prevent collisions. Their robotic shopping cart also follows users to transport goods. Up to now, only a simulation study was conducted, although the authors plan to implement the robot in a real life shopping environment.

Gai et al. [10] used the commercial Kinect sensor to enable their robotic shopping cart to execute gesture commands (e.g. a specific arm gesture of the shopper makes the robotic cart go ahead or draw back). In an experimental laboratory study with one test person the execution of different gesture commands revealed high accuracy (>93 %) and low reaction times (<0.5 s).

Vladimir Kulyukin and his work group at Utah State University developed the robot shopping cart *RoboCart* [13,36]. RoboCart is a system designed for facilitating shopping in a supermarket for visually impaired people. The robot autonomously navigates through the market, using radio-frequency identification (RFID) tags on shelves for orientation and localization. It guides customers to product locations and transports the shopping with its integrated shopping basket [34,35]. The authors also tested this robot in user studies with visually impaired people and used their comments as feedback for further development [12]. However, explicit usability tests and an analysis of robot acceptability based on an established theoretical model were not conducted.

At FZI Forschungszentrum Informatik (Research Center for Information Technology) the *Interactive Behavior Operated Trolley* (InBOT) was developed in the context of the FP6-EU project *CommRob*. There are four modes of controlling InBOT [14]: (i) Haptic Steering Mode, when the robot can be steered just as an ordinary shopping cart by the force sensitive haptic handle; (ii) Following Mode, when the robotic shopping cart follows the user in a pre-defined distance, observing the user continuously via the vision system; (iii) Guiding Mode, when the robotic shopping cart guides the user to a predefined target product or according to a shopping list, and (iv) Autonomous Mode, when the robotic shopping cart acts only if explicitly commanded by the user and once commanded performs the task independently. Test persons taking part in the usability and human-robot interaction evaluation study stated that they found it easy to get used to controlling the robot and rated the guiding functionality as useful [15].

3.2 Non-humanoid Shopping Robots

In contrast to the robotically enhanced conventional shopping carts, the shopping robot prototypes in this group possess their own designed bodies. These usually look very mechanical (non-humanoid).

For example, Tokura and co-workers use a system of two robots acting together [44] to support a purchase. This particular system consists of a guide robot, which is equipped with laser range finders that enable it to identify potential customers and offer its services as a shopping assistant. The customers can interact with the guide robot via a touch panel. The second so-called transport robot carries a shopping basket and is ready to autonomously interact with the guide robot. If the implemented transport service is selected, both robots follow the users as they move through the market. As soon as shoppers choose a product, the transport robot moves ahead so that selected items can be stashed. After placing the items, the users are asked by the guide robot if they wish to continue shopping or if they want to go to the cash desk. When the shopping is finished, both robots accompany the costumer to the cash desk [67]. Successful function tests were conducted, data on user experience, usability, and robot acceptability are lacking at this point.

Tetsuo Tomizawa and his team at the National Institute of Advanced Industrial Science and Technology (AIST) in Japan developed a shopping robot to buy groceries (especially fruit and vegetables) by controlling the robot remotely via a web interface [68]. The mobile robot is positioned in a supermarket, has a specific grappler and is connected with users via the Internet. The grappler enables remote users to pick fresh fruit and vegetables without damage and put them into a shopping cart. Further, the robot sends pictures of the goods to its users, allowing them to choose a specific item [69]. Although a function test was conducted in the field, no information on usability or acceptability of the system has been provided so far.

A research group from the Center of Artificial Intelligence at Universidad Veracruzana in Mexico [11,43] developed a non-humanoid autonomous mobile shopping robot (using the platform Pioneer 3DX from Adept Mobile Robots) that operates together with the user's mobile device (smartphone or tablet PC) and the supermarket's technological infrastructure (database servers, Wi-Fi, Bluetooth access points). The system assists shoppers by managing their shopping lists, providing product information, guiding through the supermarket and transporting the groceries. A field test with $n = 51$ end users was conducted. Both female and male shoppers gave the "robot interaction easiness" and the "overall usefulness" of the system positive ratings.

Vladimir Kulyukin and his team developed a shopping robot for visually impaired people. The Robotic Guide *dog* (RG) can be instructed by voice entry to guide visually impaired shoppers through a supermarket. The robot has a mechanical, not dog-like form, but fulfills the function of a guide dog. The implemented voice output tells users when the target is approached. By means of a bar code scanner the clients can check if they have reached for the correct article. Product information can be read aloud via voice output [35]. The robot was tested with blind users who were able to successfully arrive at the target destination. However, participants noted in their exit interviews that human-robot interaction features (for example speech recognition) was insufficient. Still, they were comfortable with the idea to use this robot in addition to white canes and guide dogs [35].

3.3 Humanoid Shopping Robots

Finally, several shopping robot prototypes are presented that possess a human-like corpus with a head (including a face) and a body (sometimes including arms, but usually no legs).

The working group surrounding Chandan Datta and Ritukar Vijay developed *Neel*, a wheeled mobile robot with indoor autonomous navigation and conversation abilities

[8]. The core idea is to enable users to interact with Neel by synergizing interaction data available from direct human-robot interaction on-site and online in the social web. The robot can store interactions on-site in its episodic memory and display them while users use the MyNeel website or mobile app. Website and app are synced in real-time with the robot's social database and use a recommendation engine. This enables the robot to suggest products and special deals in the mall. First observation studies however revealed that users were reluctant to use the robot for longer durations or to register with the robot [7]. Recent studies showed higher user acceptability rates (40 % of robot users had repeated interactions with the system) after the interaction workflow had been improved and voice/video communication between the shopping mall's customers and retailers had been implemented on the robot [71].

The research group of Koji Kamei of ATR Intelligent Robotics and Communication Laboratories in Kyoto presented a platform for ubiquitous networked robots in a laboratory shop environment [25]. Main tasks for the robots are navigation of customers and recommendation services based on prior purchases. The authors tested two scenarios: first, navigation and recommendation inside a shop [26] and second, cooperative navigation between robots, both outside and inside the shop [30]. In the shop environment, three communication robots were positioned on shelves. When a customer approached a shelf, the robot greeted the customer, introduced itself and the categories of items on the shelf. After the initial presentation, it recommended a specific item or guided the customer to another shelf, according to the customer's purchasing behavior observed by sensors before arriving at the specific shelf. Purchase behavior was observed by the spatial points the customer visited, items he or she gazed at, and items the customers picked up, using a laser range finder system, cameras, and RFID tags attached to products. The second, cooperative scenario addressed potential customers who were outside of the shop in the mall's corridor. The greeting robot's task outside the shop was to ask the customers for their shopping intentions. After recommending the customer to the shop, the greeting robot tells him or her that it will notify the other robots in the shop of the customer's visit. The greeting robot is controlled by a human operator. Customers' impressions of the shop, the sensor environments, and the robots were collected by a user survey conducted after the experiment. The shopping robots' reliability was rated high in both scenarios. Impressions of friendliness, however, were increased in the cooperative scenario.

The work group surrounding Takayuki Kanda of Advanced Telecommunications Research Institute International (ATR) in Kyoto invented *Robovie* [29,63], one of the most researched shopping robots. The focus in the development of this humanoid robot is on natural and emotional human-robot interaction. One of its tasks is to point users to specific places within a shopping mall. This is not realized by a guidance tour, instead the robot describes the way via acoustic speech output and gestures. Recent studies address questions like: How can Robovie detect a suitable waiting position in the mall without disturbing customers [32]? How can Robovie give verbal directions that are easy to understand by referring to spots in the mall that are already familiar to the specific shoppers because they have visited them together with the robot [45]? How can Robovie perform pointing gestures that are both easy to understand and polite, especially when pointing to people [39]? Should Robovie provide shopping recommendations in a whispering voice to increase the sense of confidentiality and intimacy [48]?

Robovie was also investigated in terms of user aspects [27,28,62] and application context like recommendations and advertisting in a shopping mall [49,64]. In 2007, a five-week field trial in a shopping mall was conducted with $n = 332$ pre-selected participants that the robot could identify. It turned out that 162 of the recruited participants interacted with the robot multiple times, 170 only one time, and 37 not at all. The evaluation survey (e-mail questionnaire, $n = 235$) revealed that participants indicated a high usage intention of the robot, that they were interested in the robot, and that they experienced it as familiar and intelligent. In addition, the test subjects' statements were positive regarding the directions they received from Robovie: These were evaluated as correct and surprisingly detailed. The information presented by the robot to the testers was also rated as being rather useful and interesting. Additionally, almost half of the respondents visited a particular store on basis of its instructions, and more than a quarter of respondents actually purchased a product advertised by the robot. Furthermore, Robovie was evaluated positively in terms of habituation to the system. Finally, compared to an information display Robovie was also assessed more positively. The subjects reported the information provided by the robot was more useful, more interesting, and led to more visits to retail stores and purchases there.

Also, Robovie as a humanoid shopping robot was compared to a simple robotic shopping cart in a field trial with elderly people shopping in a supermarket [23], using a Wizard-of-Oz technique to provide the service. The study served to clarify design issues for assistive robots, to answer the questions whether robots should converse with people beyond task fulfillment comments and whether people prefer a communication-oriented (humanoid) or function-oriented (cart robot) design. The field experiment with $N = 24$ elderly people showed that robots communicating and having a humanoid design increased intention to use, perceived enjoyment, but not perceived ease of use.

Fig. 1 Interactive mobile shopping robot TOOMAS based on a SCITOS A5 (developed by MetraLabs GmbH Ilmenau, Germany) with its main equipment for environment perception, navigation, and human-robot interaction

3.4 Conclusion

As related works show, a multitude of shopping robots has been developed up to now. Although in the last years considerable effort was made by researchers not only in technical robot development but also in conducting user studies, detailed and recent data on shopping robot usability and acceptability are scarce. Often, shopping robots are tested in the laboratory using very small samples, sometimes even single testers. This paper aims at closing this research gap by providing data from three separate empirical field studies with almost 350 test persons (customers of a home improvement store) of different demographics.

4 The Shopping Robot TOOMAS

The humanoid shopping robot TOOMAS was developed over the course of several years (from 2005 until 2009). Selection and implementation of technological features were based on functional and performance tests. In the following, a brief summary will be given on the technological approaches followed in the robot development project (Sect. 4.1), before derived requirements (Sect. 4.2) and the final implementation (see Fig. 1; Sect. 4.3) are presented. Afterwards, a brief overview of previous project publications on TOOMAS dealing with user evaluation studies will be given (Sect. 4.4), highlighting the original contribution of this paper.

4.1 Summary of Methodological Approaches

The simultaneous localization and mapping (SLAM) approach followed in this project used for autonomous navigation was based on a multi-modal Monte–Carlo localization [4,19], used memory-efficient grid maps for SLAM using low-resolution sonar range sensors [59] and was developed into a sensor-independent Map Match SLAM for visual mapping [60]. The robot control architecture consists of the hardware layer (enclosing sensors and actuators, the operating system, and the low-level interface to the hardware), the skill layer (including modules for collision avoidance, localization and navigation, speech recognition, speech synthesis, people tracking and so on), and the application layer (providing elements that are required for a specific application of a mobile interactive robot; [4,17]). Further, a probabilistic approach and a Bayesian update scheme have been developed for multi-modal user detection and tracking [47] for human-robot interaction, where data of various sensory systems (range measurements from occupancy gridmaps, leg detection laser scans, skin color detection, distance estimation of moving objects, detection of face images) is merged asynchronously.

4.2 Shopping Robot Requirements for TOOMAS

Specific challenges and requirements were identified in a user-centered design and development process by discussion with store owners and managers, experiences from test bed evaluations, and surveys of market staff and customers [17]. The requirements presented in the following section address general problems to be solved in shopping robot development aimed at implementation into practice.

4.2.1 Navigation and Integration

Complexity of the Environment TOOMAS was implemented in a home improvement store, which is a very complex, maze-like indoor environment. It consists of parallel, long hallways separated by shelves, and corridors connected to a network of alleys, to the main store entrance and to the checkout counters (see Fig. 2). Typically, the area has a size of 5,000 to 15,000 square meters. The environment is highly dynamic due to moving people (customers, staff) or other moving objects like shopping carts. Also, it is a highly evolutive environment regarding the filling of the shelves, or the placement of special offers at the head sides of the shelves (product heaps, sales, etc.).

Plug and Play Solutions Versus Installations The embedding of mobile service systems in a market should be feasible without cost- and time-expensive modifications of the market or its technical infrastructure. Therefore, store own-

Fig. 2 Home improvement store environment where field tests were conducted

ers usually do not favor retrofitting the whole operation area with a dense net of RFID tags or with Indoor-GPS techniques based on laser or other active components to allow a robust navigation. The best solution we identified is a simple plug & play solution using only the robot's on-board sensors and advanced navigation techniques. Further, a simple integration in existing infrastructure (like Wi-Fi, merchandize management system, etc.) is required. This serves to ensure high flexibility of the robotic solution with regard to reconfigurations of the store and changes of tasks.

Moreover, the question of how to accomplish an automatic labeling of the navigation map with the positions of all products is a key problem: a manual labeling of all article locations is impossible, since for large hardware stores, up to 60,000 different products have to be placed in the navigation map. We propose an automatic mapping between this map and the places of the articles within the shelves (stored in the merchandize management system of the store).

4.2.2 Requirements to Human–Robot Interaction

Speech-Based Dialog and Article Search Speech recognition is significantly more complex than other typical human-computer interaction scenarios due to (i) the extreme background noise in home stores, (ii) distracting public address announcements, (iii) the large diversity of descriptions for the same article, and (iv) the use of common speech and dialects, unclear articulation, and pronunciation. Therefore, it proved to be unrealistic to use speech recognition in order to identify desired articles. Implementing a pure menu-based selection on the touch screen using established methods of keyword or product line search was a solution to this problem.

Getting and Staying in Contact with Customers Establishing contact and interactive dialogs with customers has to be as intuitive as possible, avoiding prior briefing before using a shopping robot. As shopping robots to date still present a novelty, customers might be restrained or even anxious if confronted with this kind of technology. During test bed evaluations, we noticed that in 90 % of all interactions customers waited in the vicinity of the robot to be approached [17]. This fact is of immediate importance for the dialog design, particularly for the question of how the robot can take the initiative while getting in contact with customers. Moreover, RFID proved as not practicable for user detection and tracking, because most customers are not willing to wear RFID tags. Speech-based user tracking is also unsuitable because single users typically do not speak along their paths, and background noise is significantly higher than the voice of a person from a distance of several meters. However, vision in combination with distance measuring sensors proved to be suitable for robust user detection and tracking.

4.3 Robot Equipment

The *service robot platform SCITOS A5* is a joint development of the Neuroinformatics and Cognitive Robotics Lab at the Technische Universität Ilmenau with the company MetraLabs GmbH Ilmenau [17,18]. It is used under the name *TOOMAS* (based on the name of the home improvement store chain) as a humanoid and mobile shopping robot for shopping assistance. To date a total of ten service robots have been produced since April 2008 and have been employed in three stores of the toom home improvement store chain. In the following, we describe robot equipment at the end of the development process. With a height of 1.5 m the robot is comparable to the size of a 14 year-old child. Its size is optimized for a friendly appearance and an ergonomic operation. The drive system of the robot consists of a differential drive and a caster on the rear. This gives TOOMAS a good maneuverability and stability in spite of its height and weight of 75 kg, and allows a maximum driving speed of up to 1.0 m/s.

4.3.1 Sensor Equipment

For navigation, human-robot interaction, and safety, the system is equipped with various sensor systems. Firstly, there is an omnidirectional camera mounted on the top of the head. Due to the integrated hardware transformation, the camera delivers both a panoramic image (720×150 pixels) and a high resolution frontal image (720×362 pixels), which can be panned around 360°. Secondly, the robot is equipped with a set of 24 sonar sensors at the bottom, which are used for obstacle detection, map building, localization and person tracking. They cover the whole 360° space around the robot. Thirdly, a laser range finder (SICK S300) was added and mounted in front direction at a height of 35cm, as required by safety regulations of the German Technical Inspection Agency (TÜV).

4.3.2 Human–Robot Interaction

For interaction with the customers, TOOMAS is equipped with an integrated touch display, a sound system, and a 6 DOF robot head. The head turns, so that the robot's face is always looking at the customers, for example during the guidance tour. The robot was equipped with a face in order to ensure a friendly and familiar appearance (see also [41]). Eyes and eye lids are movable. Initially, the robot blinked during oral interactions with customers. However, when it was used often, blinking drained the battery very fast, leading to frequent recharging. Therefore, that feature was turned off.

The touch screen is the central human-robot interface. Customers can use it for controlling the robot, searching for articles, pricing information, starting the guidance tour, and stopping interactions with the robot.

A set of stereo loudspeakers and two omnidirectional condenser microphones are integrated in the screen device. During the first developmental stage, voice output was integrated to guide customers through the search process (for example explaining keyword search and product line search options, or giving hints on how to proceed in the search process). Several social dialogues were also integrated in order to make human-robot interaction as intuitively as possible. When approaching potential users, the robot played a greeting message: "My name is TOOMAS, can I help you?". During the guidance tour, the robot invited participants to come closer when they left the tracking range ("I can't see you anymore, would you please come closer?"). When the article location was reached, the robot told the customers that they arrived at their destination and asked whether further assistance was needed. When customers chose to end the interaction, the robot thanked the users and said goodbye.

4.3.3 Hardware

SCITOS A5 is controlled by an embedded PC with an Intel Core 2 Duo processor and a multitude of small hardware units that monitor several functions of the robot [15]. The hierarchical energy-saving concept in conjunction with the energy-saving units enables a long run-time. Based on two lead-acid gel batteries with an overall charge of 38 Ah, a SCITOS A5 autonomously operates about 8–12 h until it needs recharging. Easily connected to main supply or to its self-charging station by its integrated charging system, TOOMAS can be recharged in about 10–12 h. The safety system involves a closed bumper with tactile sensors for detection of possible collisions. In combination with additional sensors, like the vision system, the laser range finder, and the sonar sensors, SCITOS has a safety approval of the German TÜV that was given after a number of challenging safety tests.

4.3.4 Integration

TOOMAS gets required data about article information, product groups, price information, and current promotions from the market server, an off-board PC in the market which it is linked to via Wi-Fi. However, it does not rely on this connection, as it is always running from its on-board computer in fully autonomous mode. Other PCs in the market can be connected to the market server to be used as info and video-link terminals for the staff. This allows to display the current positions and statuses of all robots operating within the store.

4.3.5 Service Functionality

The task of the shopping robot TOOMAS is to accompany and support the customers during their shopping:

– *Contacting* In its role as a shopping assistant the robot is positioned at the entrance or patrolling the main corridors of the market. Using its built-in sensors, it detects customers, approaches them within an adequate communication distance and presents a multimedia-based greeting message to show its function as a shopping aid. If the customer wishes shopping assistance, more interactions follow.
– *Search for Articles* The clients can use a touch screen attached to the hull of the robot to search for desired articles using an article search system. This article search system, which is the main human-robot interface for customers, draws on information of a database that is provided and maintained by the toom Baumarkt GmbH. Users have the choice between a product line search, in which similar items are grouped into categories, and a keyword search, which can be used to search for certain products systematically. In addition, a brochure search can be conducted.

This is only available in promotional weeks when a promotional brochure of toom Baumarkt GmbH is published. In addition to a print version that is sent to households, a digital version is integrated into the database at the current time. In addition to graphics and text on a screen, a voice output is provided through integrated stereo speakers to support the search and to facilitate the operation of the robot.

– *Location Information* After a successful search, the article location and the current location of the users are shown on a map on the display. They can now choose to navigate alone or with the robot to the displayed article location.
– *Guidance Tour* If the users decide to let the robot guide them to an item's location, TOOMAS leads them through the hardware store on the shortest calculated route. TOOMAS' speed ideally equates to normal walking pace. Using its sensors (camera and distance sensors like a laser range finder and sonar) the robot maintains contact to customers and therefore is able to react by decreasing speed or stopping when, for instance, the physical distance between robot and accompanied customer increases.
– *Purchase advice* After reaching an article's location, TOOMAS offers additional services if requested. Price information can be retrieved and consultants can be contacted via video conference. If users have no further requests, they can end the interaction with the robot by simply pushing a button using the touch screen. If this does not happen, the shopping assistant discontinues interaction by itself after a certain period of time and moves to the starting point in the hardware store.

4.4 Summary of User Evaluation Studies

Besides the technological development of the shopping robot over the course of several years, the project followed a user-centered design process with an explicit focus on user factors of different target groups. Our previous publications dealt with ergonomics of the user interface and perceived opportunities and challenges of shopping robots [53]. Due to the findings of the user study, voice output was implemented, the graphical interface was redesigned and the database of the store's articles was revised (for more detail, see Sect. 5.2). Findings also showed that the implemented keyword search was more effective than the product line search and was preferred by most participants.

Further, a comprehensive evaluation of the robot's suitability to its tasks (for example voice output, extensiveness of interactions, and suitability of its movements and mobility to user needs) was conducted including comparisons between shopping experiences with or without robot assistance [42]. Details are presented in this paper, with an additional analysis of user experience of different sub-target groups (see Sect. 6.2).

Additionally, robot acceptability was compared between robot users, non users, and aborters [54]. Findings showed that users and aborters alike stated general interest in shopping robots and reported looking for a specific item as the key motivation to use the robot. Whereas users followed the robot until it guided them to the article location, aborters mostly stopped using the robot when the article was not found in the database. This finding highlights the importance of a functional search system. Non-users mostly said that they already knew where to find the article they wanted to buy or came to the home improvement store in order to browse.

We also explored subjective perceptions of robot acceptability and its determinants in the last development stage (when the robot was nearly market-ready; [55]). Findings of our user study showed that intention to use the robot was rather high and influenced by a positive attitude towards the technology, meaning that the robot was perceived as easily operable and as a helpful shopping assistant.

Lastly, robot personality and its influence on robot acceptability was analyzed by comparing a decidedly extraverted version of TOOMAS to the conventional version [41]. The extraverted robot had a nose, a smiling mouth and lashes. The voice varied in pitch, verbal expressions were more positive, accompanied by frequent nodding, and more frequent winking. Findings revealed that the extraverted robot was clearly preferred by the customers. Especially the nonverbal communication features seem to have enriched human-robot interaction and increased acceptability.

In contrast to those publications, in this paper we present previously non-published and more detailed results on robot usability and acceptability differentiating between different sub-target groups based on robot user demographics (gender, age, educational background and computer skills levels).

5 Developmental Stage 1: Article Search System

In this section, we present a formative evaluation study serving to optimize the usability of the article search system of our shopping robot TOOMAS. We followed a user-centered design process for the development of the search engine and the graphical user interface, as this serves as the most important human-robot interface for the robot at hand. If human-robot communication fails at this point, no other interactions with the robot will take place. Usability requirements for search engines and user interfaces are particularly high, as clients from different population groups are to use the system freely and without previous learning [3].

A precise usability evaluation was necessary in order to know to what extent the requirements for a user-friendly human-robot interaction were fulfilled (effectiveness, efficiency, and satisfaction according to ISO norm ISO 9241-11). Further, we explored if sociodemographics (gender, age, educational background) and computer skills had an influence on usability. The following research questions were to be answered:

- RQ1: Are customers able to retrieve sought-after articles using the robot's article search system (effectiveness)?
- RQ2: Are customers able to retrieve sought-after articles in a short time-frame using the robot's article search system (efficiency)?
- RQ3: Is using the robot's article search system agreeable and satisfactory for customers (satisfaction)?
- RQ4: Do gender, age, educational background, and computer skills have an effect on the usability of the robot's article search system?

We also explored robot acceptability by store customers, using intention to use as an indicator (based on [70]). As the robot was seldom available, customers were presented a poster introducing the robot and its features. The following research question was formulated:

- RQ5: Are customers willing to use the robot as a shopping assistant in the future? Do gender, age, educational background, and computer skills have an effect on future intention to use the shopping robot?

Also, we followed a formative evaluation procedure [61], which produces ongoing intermediate assessments of evaluation objects. Those assessments are used for an iterative optimization process. The main goal is to identify concrete usability problems and solve them step by step along the process.

5.1 Method

5.1.1 Design

A non-experimental cross-sectional formative evaluation study was conducted with three consecutive phases. Every phase consisted of data collection in the field and statistical data analysis. Results gained in every data collection phase were analyzed by social scientists, discussed with robot developers, and included in a usability engineering process for the user interface and the search engine installed. During the three data collection phases (taking part between August 2005 and October 2005) usability criteria were tested with different samples, as it could not be guaranteed that the same walk-in customers would be available for every measurement point during the study.

Oral interviews were used for subjective usability measures like satisfaction, and overt observation for technical malfunctions and objective usability measures like effectiveness and efficiency. Field trials were conducted in a home improvement store of toom Baumarkt Gmbh in Erfurt in Germany.

5.1.2 Participants

In order to follow a user-centered process, it was necessary to test the shopping robot with potential end users, in our case customers of home improvement stores. Due to ethical issues, all participants were volunteers, in all cases walk-in shoppers on the days the study was conducted who used the shopping robot for the very first time. After their consent, they were included in the ad-hoc sample. Our aim was to optimize the robot in such a way that novice users are able to successfully interact with the robot. Customers who already had experience with the robot were excluded from the sample.

In every one of the three subsequent data collection phases of the first formative evaluation study $n = 70$ participants took part, resulting in a final convenience sample of $N = 210$ subjects. Details on the sociodemographic and control variables are presented in Table 1. In order to check for correlations between those vari-

Table 1 Descriptive statistics of sociodemographic and control variables for the total sample of study 1

Variable	n	%
Gender		
Men	125	59.5
Women	85	40.5
Total	210	100
Age group		
16–29	36	17.1
30–39	36	17.1
40–49	55	26.2
50–64	68	32.4
65 & older	15	7.1
Total	210	100
Educational level		
Secondary modern school (years 5-9)	34	16.2
Middle school	106	50.5
High school	29	13.8
University	41	19.5
Total	210	100
Computer skills level		
Beginners	50	23.8
Average users	100	47.6
Advanced users	53	25.2
Professionals	7	3.3
Total	210	100

ables, we computed exact Fisher tests (as some of the cells had an $n < 5$), showing no significant correlations between gender, age, educational level, and computer skills respectively.

5.1.3 Materials

To measure subjective usability criteria and future intention to use the robot, structured interviews were conducted, using five-point Likert scales (from 1 = do not agree to 5 = do fully agree).

Further, sociodemographic variables and computer skills were surveyed. Considering prompting and probing, oral interviews are more effective than written questionnaires as they dealt with an unfamiliar topic (shopping robots in the home improvement store setting [5]).

Satisfaction (as one of the established usability criteria according to ISO 9241-11) was measured by a scale consisting of the following two items: "Were you satisfied with the search result?" and "How would you rate the system in general?". Future intention to use was measured before and after participants used the system for an article search by the following item: "Would you use the robot yourself?" Although the robot in itself was not tested, we used this item in order to ensure compatibility with later user studies testing TOOMAS.

Moreover, two observers recorded objective usability criteria, for example whether customers were able to find sought-after articles by using the search engine (articles found or not, effectiveness) and search duration (search time taken by means of a stop watch, efficiency). Due to the obvious nature of those criteria, inter-rater reliability was practically 100 %.

5.1.4 Procedure

After being recruited by a professional field interviewer (people especially trained for recruitment) at the store entrance, subjects received basic information on the study and on data protection. After informed consent was given, sociodemographic variables were collected. In the next step, the robot was presented on a poster and its features were explained to the participants. In addition, for usability testing, subjects were asked to use the robot's search engine in order to search for one article of their own choice and its location. System usage was observed to record usability problems and handling difficulties. After testing the search engine, users were interviewed on their experience with the robot and intention to use the robot in the future. The user study was conducted at a stationary terminal and not with the mobile robot as it was seldom available due to intensive technical system revisions.

5.2 Results

Firstly, usability engineering of the user interface will be presented (Sect. 5.2.1). Afterwards, findings of the usability evaluation over the course of the first study's evaluation phases as well as according to sociodemographic and control variables will be given (Sect. 5.2.2) before corresponding results on robot acceptability will be presented (Sect. 5.2.3).

5.2.1 Usability-Engineering of User Interface

After every data collection phase, interviews and observation protocols were analyzed and system developers were informed about identified usability problems. Revisions of the user interface and the search engine are presented in the following, classified by relevance.

Voice Output The most important revision was the implementation of a voice output, guiding participants through the search process and giving hints for using the article search system. During the first test phase, some subjects hesitated after having inserted a search keyword. Several product lines were presented in the interface, possibly containing the article searched for. Most subjects did not realize that they were able to select a specific product line. In this case, from the second phase on, voice output gave feedback after seven seconds, telling the user to choose a certain category. Further, voice output explained alternative ways of searching when articles were not found.

Graphical Redesign of User Interface Due to the findings from the first data collection phase, the user interface was redesigned before the second test phase as follows (Fig. 3 shows the interface at the beginning and the end of this first formative evaluation study).

Most subjects in the first phase hesitated because they did not realize that presented product lines were clickable and looked for an 'enter' button. To handle this problem, buttons were designed in a more three-dimensional way. In the original version, the back button was placed on the upper left side of the interface, as is common in web browsers. However, some users did not detect it. During redesign, it was placed in the lower right of the interface more closely to other control elements. Instead of the back button, a main menu button was inserted at the upper left side, leading to the option between product line search and keyword search. Most of the handling problems were addressed by the redesign and decreased significantly over the three testing phases (see Table 2).

Revision of Database The robot's article search engine was based on a database originally constructed for stocktaking

Fig. 3 Exemplary presentation of user interface redesign (left: beginning of study, right: end of study)

Table 2 Absolute frequencies of handling problems (obtained by observation) compared for data collection phases

	Product lines not clicked on	Search for enter button	Back button not detected
Phase 1	17	16	9
Phase 2	11	10	3
Phase 3	1	4	0
χ^2	16	8.4	11.1
p	<.05	<.05	<.05
n	209	210	209
Cramer's V	.28	.20	.23

$df = 2$, different n due to missing data

by toom Baumarkt GmbH. This database consists of more than 60,000 different products, sometimes with several variations (for example various sizes of screws), which leads to over 80,000 articles in the database. Many terms and definitions did not conform with everyday language use. Further, not every article was included with its actual location. These problems decreased effectiveness, as some articles could not be found by using the search engine. A list of problematic key words was generated and sent to the home improvement store management. Besides, location declaration was adapted. Although the problem could not be solved completely, the rate of declared article locations increased significantly over the three data collection phases (see Table 3 for location presentation when the article was found in the database). In the first phase, 83% of search requests led to accurate article locations, in the third phase, this was the case in 98% of requests ($\chi^2 = 15.3$; $df = 2$; $p < .001$; *Cramer's V* $= .35$; $n = 126$). Whereas the redesign of the graphical user interface and voice output integration was accomplished before the second test phase started, revision of the database was ongoing.

Table 3 Article location (in percent) of articles found for three data collection phases

	Article location presented (in %)	Article location not presented (in %)	n
Phase 1	83	17	35
Phase 2	67	33	45
Phase 3	98	2	46
Total	83	17	126

5.2.2 Usability

Further, we analyzed the usability of the user interface and the search engine as well as their changes during the three test phases. Also, we examined usability with regard to sociodemographic data and control variables. First, findings for usability over the course of the three evaluation phases will be presented before detailed results for gender, age group, educational background and computer skills will be reported.

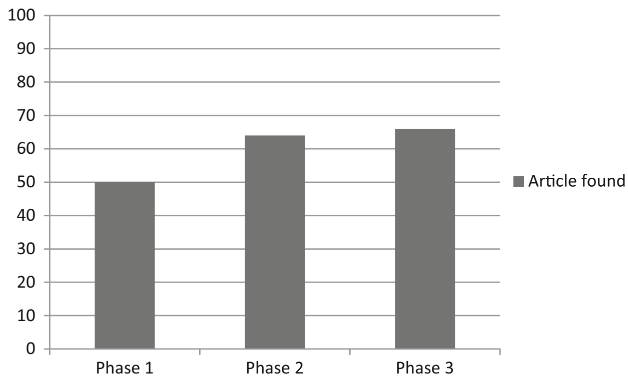

Fig. 4 Rate of successful search requests (in percent) by means of product line search and keyword search for the three data collection phases

Usability Over the Course of the Evaluation Phases Considering effectiveness, half of the search requests during the first inquiry stage were successful. After a revision of the search engine, search success increased up to 64 %. The next revision made it up to a success rate of 66 % (see Fig. 4).

This increase in effectiveness was not significant ($\chi^2 = 4.4$; $df = 2$; $p = .11$; *Cramer's* $V = .15$; $n = 210$). However, the 34% unsuccessful search requests during the last phase were not a consequence of usability problems but of aforementioned problems with the product database provided to us by the store management. A basal revision of the database would have been very cost and time consuming, due to its complexity (more than 80,000 items). On that account, it could not be put into effect at the time of research.

In regard to efficiency, a similar pattern was found: One search request using the search engine lasted on average 163 seconds ($SD = 87$s). In course of the three inquiry stages no significant decreases of search time were noticed ($F = 2.7$; $df = 2$; $p = 0.07$; *partial* $\eta^2 = 0.03$; $n = 210$), but compared to the first stage with an approximate search time of 2m54s (*SD* = 88 s), in the third stage the search process was terminated after 2 m 23 s (*SD* = 87 s) in average.

After using the article search system, customers declared their satisfaction with the article search engine ($M = 3.9$; $SD = 0.9$; five-point Likert scale from 1= very unsatisfied to 5= very satisfied), which did not vary significantly during the inquiry stages ($F = 1.6$; $df = 2$; $p = .20$; *partial* $\eta^2 = 0.15$; $n = 210$).

Usability According to Sociodemographic Variables We further analyzed usability criteria according to sociodemographic and control variables for the total sample over all three test phases of the study. Gender showed neither significant differences considering effectiveness ($\chi^2 = 0.08$; $df = 1$; $p = .77$; *Cramer's* $V = 0.02$; $n = 210$) nor satisfaction ($M_{male} = 3.92$, $SD = 0.84$, $N = 125$; $M_{female} = 3.89$, $SD = 0.85$; $N = 85$; $t = 0.30$; $d = 0.04$; $df = 208$; $p = .77$; $d = 0.04$). However, women were significantly faster during the search ($M_{male} = 175.18$s, $SD_{male} = 92.31$s, $N = 125$; $M_{female} = 144.96$s, $SD_{female} = 76.56$s, $N = 85$; $t = 2.49$; $df = 208$; $p = .014$; $d = .36$).

Table 4 Absolute frequencies of search success (articles found) according to age group

Age group	Article found	Article not found	*n*
16–29	25	11	36
30–39	19	17	36
40–49	35	20	55
50–64	43	25	68
65 & older	4	11	15
Total	126	84	210

Table 5 Means and standard deviations for search duration in seconds (efficiency) according to age group

Age group	*M*	*SD*	*n*
16–29	149.11	78.05	36
30–39	141.03	73.70	36
40–49	152.30	79.07	55
50–64	174.10	93.62	68
65 & older	237.20	101.95	15
Total	162.95	101.95	210

Age group also showed a significant effect for effectiveness (see Table 4). Although in almost every age group more articles were found than not found, for seniors being 65 years old or older, the reverse effect was shown ($\chi^2 = 9.66$; $df = 4$; $p = .05$; *Cramer's* $V = .22$; $n = 210$). In order to determine what age group led to this effect, we conducted post hoc χ^2-tests with Bonferroni adjustment to prevent type I error inflation due to multiple comparisons (adjusted alpha level = .005). No significant differences were shown. However, the standardized residual for articles not found by participants being 65 years and older was 2.0, indicating that the effect resulted from this cell.

Although age did not lead to a significant effect considering satisfaction ($F = 0.11$; $df = 4$; $p = .978$; $n = 210$, *partial* $\eta^2 = .002$), it did for efficiency ($F = 4.23$; $df = 4$; $p = .003$; $n = 210$, *partial* $\eta^2 = .076$, see Table 5). Games-Howell post hoc tests confirmed that seniors of 65 years of age and older showed significantly longer search durations than younger age groups ($p < 0.05$) with the exception of participants between 50–65 years ($p = 0.22$).

Educational background showed no significant effect for effectiveness ($\chi^2 = 2.66$; $df = 3$; $p = .45$; *Cramer's* $V = 0.11$; $n = 210$). Neither did satisfaction differ between educational levels ($F = 0.51$; $df = 3$; $p = .678$; $n =$

Table 6 Means and standard deviations for search duration in seconds (efficiency) according to educational level

Educational level	*M*	*SD*	*n*
Secondary modern school (years 5–9)	185.74	85.52	34
Middle school	170.01	87.48	106
High school	152.17	85.35	29
University	133.41	84.30	41
Total	162.95	87.36	210

Table 7 Absolute frequencies of search success (articles found) according to computer skills level

Computer skills	Article found	Article not found	Total
Beginners	22	28	50
Average users	67	33	100
Advanced users	32	21	53
Professionals	5	2	7
Total	126	84	210

Table 8 Means and standard deviations for search duration in seconds (efficiency) according to computer skills level

Computer skills	*M*	*SD*	*n*
Beginners	201.64	85.02	50
Average users	161.22	85.77	100
Advanced users	134.60	80.75	53
Professionals	125.86	87.87	7
Total	162.95	87.36	210

210, *partial* $\eta^2 = 0.007$). However, considering efficiency again, search duration significantly descreased with higher education levels (see Table 6; $F = 2.78$; $df = 3$; $p = .042$; $n = 210$; *partial* $\eta^2 = 0.039$). Games-Howell post hoc tests confirmed that university graduates were significantly faster when searching for an article than participants who graduated from secondary modern school ($p = .047$).

Computer skills showed no significant effect for effectiveness ($\chi^2 = 7.76$; $df = 3$; $p = .051$; *Cramer's V* $= 0.19$; $n = 210$). However, this effect barely missed reaching statistical significance. Beginners showed more unsuccessful searches (see Table 7).

Neither did satisfaction differ between computer skill levels ($F = 0.69$; $df = 3$; $p = .558$; $n = 210$, *partial* $\eta^2 = 0.010$). However, considering efficiency, search duration significantly descreased with higher computer skills (see Table 8; $F = 5.96$; $df = 3$; $p = .001$; $n = 210$, *partial* $\eta^2 = .080$). Games-Howell post hoc tests revealed that beginners took significantly more time to accomplish the search task than average users and advanced users ($p = .04$). However, due to the low number of professionals and therefore insufficient statistical power, there was no significant post hoc effect between beginners and professionals ($p = .22$).

5.2.3 Robot Acceptability

Additionally, we analyzed if future intention to use a shopping robot changed after participants had used the article search system developed for TOOMAS. Further, we tested the influence of the sociodemographic variables and computer skills on intention to use the robot after participants were interacting with the system. As the robot was a novelty and all subjects in the sample were first-time robot users, intention to use the robot measured before the usability test served as a baseline.

Intention to use the robot before the usability test was rather high ($M_{before} = 3.89$; $SD = 1.11$, $N = 210$) and increased slightly after using the article search system ($M_{after} = 4.00$; $SD = 1.11$; $N = 210$), although not significantly ($t = 1.74$; $df = 209$; $p = .083$; $d = 0.12$).

Neither gender ($t = 1.09$; $df = 208$; $p = .277$, $d = 0.15$), age ($F = 1.76$; $df = 4$; $p = .140$; *partial* $\eta^2 = 0.04$), educational level ($F = 0.68$; $df = 3$; $p = .563$; *partial* $\eta^2 = 0.012$), nor computer skills ($F = 1.18$; $df = 3$; $p = .318$, *partial* $\eta^2 = 0.021$) showed a statistical significant effect on intention to use the robot after the participants had used the article search system (see Table 9).

5.3 Discussion

In this study, usability engineering on basis of interviews and usability tests with customers was successful. Derived user needs were acknowledged and drawn on in iterative system optimizations: It consisted of an integration of voice output to help customers with the search process, a redesign of the graphical user interface and a revision of parts of the connected database. Handling problems decreased significantly over the three data collection phases. However, effectiveness and efficiency showed only a non-significant tendency to increase over the three data collection phases. A reason for this might be that until the end of the study, sought-after articles sometimes were not included in the article database provided by the store management.

Further, the detailed examination of usability according to sociodemographic and control variables led to some interesting results. Although effectiveness did not show significant differences between most of the sub-groups, an effect was found for age, showing that seniors being 65 years or older had a significantly lower proportion of successful searches. Maybe younger people were better able to cope with handling problems identified at this stage of development. Additionally, sub-groups showed different results for efficiency: Women, younger participants, subjects with a higher edu-

Table 9 Means and standard deviations for intention to use after testing the article search system according to gender, age, educational level, and computer skills level

Control variable	*M*	*SD*	*n*
Gender			
Men	3.94	1.14	125
Women	4.11	1.06	85
Age group			
16–29	3.89	1.09	36
30–39	4.25	0.97	36
40–49	4.00	1.02	55
50–64	3.97	1.23	68
65 & older	3.87	1.25	15
Educational level			
Secondary modern school (years 5–9)	4.29	0.94	34
Middle school	4.06	1.02	106
High school	3.62	1.29	29
University	3.90	1.26	41
Computer skills level			
Beginners	4.08	1.05	50
Average users	4.11	1.04	100
Advanced users	3.75	1.25	53
Professionals	3.86	1.21	7
Total	4.00	1.11	210

Likert scale from 1 = totally disagree to 5 = totally agree

cational background, and higher computer skills were significantly faster in their search for specific articles. Although this effect might be explained with affinity for technology and practical experience with technological systems for age, educational background and—naturally—computer skills, this is not the case for gender. It has to be noted that the investigators also were women. Therefore this might be explained by an investigator effect with female participants feeling more comfortable in an interaction with female researchers.

Despite of these drawbacks in effectiveness and efficiency, satisfaction with the system as well as intention to use the robot in the future were rather high from the beginning on and did not differ over time or according to gender, age, educational level, or computer skills level.

6 Developmental Stage 2: Adaptation to the Robot's Role and Task

During the second developmental stage, the article search system was already implemented in the shopping robot. Within this second study, developers wanted to optimize the robot and meet user requirements concerning its task and role. In addition, system acceptability of future customers was to be examined before the commercial launch of the shopping robot in home improvement stores. Therefore, a formative evaluation study was conducted in order to eliminate constraints and technical problems with the robot. The following research questions were to be answered:

- RQ1: Do embodiment, mobility, voice output, and personality of the shopping robot meet the needs and requirements of users with different demographics?
- RQ2: Are customers willing to use the robot as a shopping assistant in the future? Do gender, age, educational background, and computer skills have an effect on future intention to use the shopping robot?

6.1 Method

6.1.1 Design

A non-experimental cross-sectional field study was conducted. Oral interviews were conducted to obtain user evaluation on usability criteria as well as acceptability. Overt observation was used to obtain data on technical malfunctions.

6.1.2 Participants

Subjects were recruited by a professional interviewer in the entrance area of the home improvement store. Walk-in shoppers on the days the study was conducted were asked to participate. After giving their informed consent, they were included in the ad-hoc sample, resulting in a final sample of $N = 39$ subjects. Details on the sociodemographic and control variables are presented in Table 10. In order to check for correlations between those variables, we computed exact Fisher tests (as some of the cells had an $n < 5$), showing no significant correlations between gender, age, educational level, and computer skills respectively.

6.1.3 Materials

Evaluation criteria like understandability of voice output and adequate embodiment of the robot were operationalized with five-point Likert scales (from 1 = do not agree to 5 = do fully agree). Sociodemographic variables were also obtained, using oral interviews. Technical malfunctions were collected via observation and log files, for instance whether human-robot interaction was disturbed, e.g. when the robot lost contact with the participant during a guidance tour.

Table 10 Descriptive statistics of sociodemographic and control variables for the total sample of study 2

Variable	*n*	%
Gender		
Men	28	71.8
Women	11	28.2
Total	39	100
Age group		
16–29	3	7.7
30–39	4	10.3
40–49	9	23.1
50–64	19	48.7
65 & older	4	10.3
Total	39	100
Educational level		
Secondary modern school (years 5-9)	5	12.8
Middle school	20	51.3
High school	10	25.6
University	4	10.3
Total	39	100
Computer skills level		
Beginners	9	23.1
Average users	20	51.3
Advanced users	8	20.5
Professionals	2	5.1
Total	39	100

6.1.4 Procedure

Field trials were conducted in a home improvement store of toom Baumarkt GmbH in Erfurt in Germany in December 2006.

After being recruited by a professional interviewer, subjects received basic information on the study and on data protection. After consent, sociodemographic variables were collected. During subsequent usability tests, subjects were asked to use the robot to search for an article and its location. System usage was observed to record usability problems. After testing the shopping robot, users were interviewed about their experience with the robot, including robot acceptability.

6.2 Results

6.2.1 Technical Malfunctions of the Shopping Robot

Observation and log files of the human-robot interactions were used to keep record of technical malfunctions and to optimize the robot. The noticed technical difficulties are discussed in order of their relevance.

Non-recognition of Customers At the time of the study, one elementary problem was that the robot began to move away when interested customers were moving towards it (11 %, $n = 4$). The main reason for this problem was the not perfectly working person-tracking function. This defect also affected the guidance tour to the article positions in the home improvement store. Although customers were following the shopping robot during the tour, the robot was partly unable to track them. The statement "I cannot see you anymore" led to customer's confusion and averted a fluent tour to the location of the article.

Orientation Problems Malfunctions of the robot's detection system led to orientation problems and 'accidents' ($n = 6$). For example, the robot cut the curve too closely and hit objects like racks or shopping trolleys.

Voice Output Integrated voice output was installed to help customers to use the robot. However, TOOMAS sometimes quitted in the middle of the sentence. This problem was mainly notable when complex situations were at hand. For example when the robot wanted to turn around, tried to warn the customer at the same time of its movement, and then indicated the customers to follow.

The demonstrated problems concern important criteria for a successful human-robot interaction. Non-recognition of customers, orientation problems and problems with voice output are indicators for an insufficient adaptation of the robot to its environment and tasks. The results about the technical dysfunctions were handed to the developer team who optimized the system after the study.

6.2.2 Adaptation to Role and Task

Relevant for adaptation of the robot to its main task and role in the home improvement store environment is an adequate appearance and embodiment. During the first inquiry stage, the robot still looked very technical (see Fig. 5), because the final casing was not integrated yet. Accordingly, only 64 % ($n = 25$) of respondents liked the design of the service-robot. The majority of the 14 persons who disliked the appearance (22 %) said that they missed the envelope of the service-robot. Exact Fisher tests (as some of the cells had an n < 5) did not show any significant correlations between participants' rating of the design on one the hand, and gender, age, educational background, and computer skills on the other hand.

In reference to the integrated voice output, the study showed that enunciation was very comfortable for customers. With the help of a five-point Likert scale, the enunciation was mainly rated as not complicated ($M = 1.4$; $SD = 0.90$) and well articulated ($M = 4.8$; $SD = 0.70$). Speech speed ($M = 3.0$; $SD = 0.20$) and speech volume ($M = 3.0$;

Fig. 5 Shopping Robot TOOMAS during and after the second study (from *left* to *right*)

Table 11 Evaluation of enunciation as complicated according to computer skills level

Computer skills	*n*	*Mean rank*
Beginners	9	27.17
Average users	20	18.77
Advanced users	8	16.0
Professionals	2	16.0

Table 12 Evaluation of the robot's approaching behavior

	M	*SD*	*n*
Friendly	4.7	0.5	39
Comfortable	4.3	0.9	38
Clumsy	2.2	1.3	39
Intrusive	1.4	0.9	39

Likert scale from 1 = totally disagree to 5 = totally agree

Table 13 Evaluation of friendliness of the robot's approaching behavior according to computer skills level

Computer skills	*n*	*Mean Rank*
Beginners	9	22.39
Average users	20	21.65
Advanced users	8	16.81
Professionals	2	5.50

$SD = 0.40$) were seen as adequate and matched the environment. Sociodemographic variables showed no influence on voice output ratings, with one exception: participants with lower computer skills rated enunciation as more complicated ($\chi^2 = 10.15$; $df = 3$; $p = .005$; $n = 39$; see Table 11).

Extensiveness of human–robot-interactions was rated as complete enough (by 95 % of subjects), with exact Fisher tests showing no significant correlations between sociodemographic variables and computer skills level.

Considering adaptation of the robot's *Movement and Mobility* to user requirements it was revealed that 59 % of subjects rated driving speed as comfortable. However, the robot was too slow for 33 % of customers and too fast for another 7 %. Again, no significant correlations with sociodemographic variables and computer skills appeared.

Regarding social behavior, TOOMAS appealed to a rather high degree. Human interaction partners rated the robot approaching them as comfortable and friendly, and not as intrusive or clumsy (see Table 12).

However, in contrast to the sociodemographic variables, computer skills again showed a significant effect concerning friendliness of the approach: beginners and participants with average computer skills rated the robot's approach as friendlier than advanced and professional computer users ($\chi^2 = 8.71$; $df = 3$; $p = .012$; $n = 39$; see Table 13).

6.2.3 Intention to Use

Most of the subjects (87 %) reported intention to use the robot again, especially to search for specific articles ($M = 4.8$; $SD = 0.40$), partially for getting pricing information ($M = 3.6$; $SD = 1.5$), but less for detailed product information ($M = 2.8$; $SD = 1.6$). Sociodemographic variables and computer skills showed no influence on intention to use the robot.

6.3 Discussion

In the presented second study, the robot's optimization to its role and task (embodiment, voice output style, extensiveness of interactions, mobility, and social behavior) was analyzed by means of user tests. The findings provided valuable hints for continuous system optimizations. Several proposals were implemented during the course of the conducted tests. However, optimizations met their limits: again, certain articles were not found in the database. Database revision could still only take place continuously and in parts. In summary, robot usability with regard to its interaction with human partners was evaluated as relatively positive. Mainly, sociodemographic variables showed no correlation with usability measures. However, computer skills revealed two effects: participants with lower computer skills rated the robot's enunciation as more complicated, but experienced the robot as friendlier when it approached them. The first finding highlights the need to simplify content of voice output. The second finding could be interpreted as a novelty effect: participants with less experience in handling computers might be more impressed by the integrated social interactions, whereas subjects with more computer experience might also have higher expectations regarding human-robot interaction. Finally, acceptability was rather high: Most of the participants reported intention to use the shopping robot in the future, especially for article search.

7 Developmental Stage 3: Comparison of Robot-Assisted and Conventional Shopping

In 2009, at the end of the pilot phase, a summative evaluation study of TOOMAS was conducted. It had to be examined if effectiveness, efficiency and customer satisfaction requirements (DIN EN ISO 9241-11) were met. Fulfillment of these requirements was compared to shopping without robot assistance. The results gained in this study served decision-making on potential continuation of the project. The following research questions had to be answered:

– RQ1: Is shopping with robot assistance more effective than conventional shopping?
– RQ2: Is shopping with robot assistance more efficient than conventional shopping?
– RQ3: Is shopping with robot assistance more satisfactory for customers than conventional shopping?

Further, intention to use was again measured as an indicator of acceptability.

Table 14 Descriptive statistics of sociodemographic and control variables for the total sample of study 3

Variable	*n*	%
Gender		
Men	54	57.4
Women	40	42.6
Total	94	100
Age group		
16–29	13	13.8
30–39	9	9.6
40–49	22	23.4
50–64	31	33
65 & older	19	20.2
Total	94	100
Educational level		
Secondary modern school (years 5–9)	31	33
Middle school	28	29.8
High school	26	27.7
University	9	9.6
Total	94	100
Computer skills level		
Beginners	19	20.2
Average users	44	46.8
Advanced users	23	24.5
Professionals	8	8.5
Total	94	100

7.1 Method

7.1.1 Design

A quasi-experimental between-subject field study was conducted. Participants were assigned to two conditions: shopping with robot assistance and conventional shopping, creating matched pairs using sociodemographic data and computer skills to control for personal variables between groups. Oral interviews were used to measure satisfaction and intention to use, whereas overt observation was used to gain data on effectiveness and efficiency. The field experiment was conducted in 2009 in a home improvement store of toom Baumarkt GmbH in Bergheim, Germany.

7.1.2 Participants

A total of $n = 94$ home improvement store customers took part in the study. Those were recruited in the entrance area of the home improvement store and included in the convenience sample after giving informed consent. Sociodemographic data are presented in Table 14. Participants had a mean age of $M = 49.9$ years ($SD = 14.4$) ranging from

16 to 74 years. Contrary to the previous studies, age was not measured with categorial data, but asked for directly. This served to create more accurate matched pairs. Participants were arbitrarily subdivided into experimental group (shopping with robot assistance, $n = 47$) and control group (conventional shopping without robot assistance; $n = 47$). Randomization was not feasible, due to low customer appearance on certain days. Instead, matched pairs were formed to ensure high comparability between groups. Matched pairs were formed by consecutive steps, first matching for gender, then age, eductional background, and lastly computer skills. Age difference between the participants in one pair was kept as small as possbile. For categorial data, participants from the same group (for example advanced computer skills) were matched.

7.1.3 Materials

Evaluation criteria like satisfaction with article search and intention to use the robot were operationalized with five-point Likert scales (from 1 = do not agree to 5 = do fully agree). Sociodemographic variables were also obtained, using oral interviews for data collection. Some evaluation criteria were collected via overt observation. Two observers recorded whether sought-after articles were obtained (effectiveness) and measured objective search duration (efficiency) using a stop watch.

7.1.4 Procedure

After being recruited by an investigator, subjects received basic information on the study and on data protection. Further, they were asked whether they came to the store in order to buy specific items and if they did or did not know their locations. The shopping robots should be tested with users inclined to use the system. People only strolling or looking for articles with known location would probably not profit from using the shopping assistant TOOMAS and were therefore excluded from the study.

After giving informed consent, subjects were arbitrarily assigned to either the experimental group or the control group and asked to search for the item they intended to buy. Both kinds of article searches were observed to protocol effectiveness and efficiency. After the shopping tour, users were interviewed on sociodemographics, satisfaction with their shopping activities, and intention to use the robot in the future.

7.2 Results

7.2.1 Effectiveness

With regard to effectiveness, no significant difference between retrieving an article with or without robot assistance was revealed (see Table 15; *McNemar* $\chi^2 = 1.0$; $df = 1$; $p = 1.0$; $n = 47$ measurement pairs). All in all, findings show that most participants were able to find sought-after articles with or without robot assistance respectively.

Table 15 Observed combination of pairs of measured values for robot users and robot non-users in regard to finding sought-after articles ($n = 47$ pairs of measured values)

	Robot users		
	Article found	Article not found	Sum of row
Robot non users			
Article found	30	6	36
Article not found	7	4	11
Sum of column	37	10	47

7.2.2 Efficiency

Considering efficiency, the objective search duration showed a significant difference. Searching without the shopping robot was faster ($M = 135.4$ s; $SD = 85.8$) than searching with robot assistance ($M = 180.3$ s; $SD = 82.8$; $t = 2.8$; $df = 46$; $p = .003$; $n = 94$). Effect size was medium ($d = 0.40$). Therefore, shopping with robot assistance was less efficient than conventional shopping.

7.2.3 Satisfaction and Intention to Use

Regarding customer satisfaction there was no significant difference between shopping with ($M = 4.1$; $SD = 1.0$) and without robot assistance ($M = 4.1$; $SD = 0.90$; $t = 0.30$; $df = 46$; $p = .74$; $n = 94$; $d = 0.05$).

However, participants who used the robot expressed high intention to use the robot in the future ($M = 4.11$; $SD = 1.32$; $n = 47$).

7.3 Discussion

Shopping with TOOMAS was not more effective, efficient or satisfactory than conventional shopping without the robot. This could be caused by the following reasons: First of all, residual/final bugs in the store's article database could have unnecessarily extended search duration. For example, some sought-after articles were still not included in the database. Secondly, participants used the robot for the very first time. It can be expected that operation duration will be shorter when people are more familiar with handling TOOMAS. Lastly, the robot operates more slowly when the store is busy and the aisles are full of moving customers, shopping carts and staff. Still, measured effectiveness, efficiency and satisfaction for shopping with robot assistance were not notably worse compared to conventional shopping, and future intention to use the robot was relatively high with robot users. Some non-users (53 % of subjects; $n = 29$) did not retrieve articles without help, but asked for help at the information desk either immediately (38 %; $n = 29$) or after an unsuccessful attempt to find the article on their own (21 %; $n = 29$).

8 General Discussion

The humanoid mobile shopping robot TOOMAS was developed for long-term everyday use in spacious and complex home improvement stores. Its main function is to guide customers directly to the locations of the products they are interested in. Customers meet the robot at the store entrance and interact with it using a graphical user interface on a touch-screen where they can select the sought-after articles from a database. In addition, they can retrieve price and further product information.

To ensure usability and acceptability of the shopping robot the development of TOOMAS followed a user-centered design process from the very beginning. This paper presented three consecutive evaluation studies addressing usability and acceptability that were conducted in different home improvement stores involving about 350 customers altogether as research subjects.

The *first formative evaluation study* investigated the usability of the robot's article search system. Several usability problems could be identified and eliminated leading to significantly more successful article searches. In general, participating customers perceived using the robot's article search system as satisfactory and reported high intention to use the shopping robot in the future.

The *second formative evaluation study* focused on the robot's adaptation to its main task and role. Embodiment, voice output, mobility, approaching behavior towards human interaction partners, and human-robot interaction were evaluated. Again, some usability and acceptability problems were identified and eliminated (e.g. regarding the robot's appearance). Most participants reported high intention to use the shopping robot in the future.

The first two studies revealed interesting effects regarding sociodemographic variables: In the first study, seniors of 65 years or older were less effective when using the article search system. Whereas women, younger participants, subjects with a higher educational background and more computer skills were faster during their search. In the second study, subjects with lower computer skills rated the robot's enunciation as more complicated, but experienced the robot's approaching behavior as friendlier. According to these findings the needs of older, less formally educated and less computer experienced end users should be taken into account to satisfy the whole diverse target group of home improvement store customers.

At the end of pilot phase, a *third summative evaluation* study was conducted. Effectiveness, efficiency and satisfaction with robot-assisted shopping was compared to conventional shopping without robot assistance. Further, robot acceptability was measured. While conventional shopping turned out to be faster, there were no differences in customer satisfaction between experimental and control group. Again, participants' intention to use the robot in the future was high.

Some limitations of the three presented evaluation studies need to be kept in mind. Data collection was partly conducted by overt observations and oral interviews. Therefore, social desirability and investigator effects cannot be excluded. Although our research involved an unusually large number of participants, generalizability is still limited because convenience sampling (instead of random sampling) was used. Our research subjects were first time users of the shopping robot, therefore questions regarding long-term experience and acceptability remain open. Intention to use the robot in the future was measured as an indicator of acceptability but these self-report data need to be confirmed by objective data on future use behavior. While some of the identified usability problems could be eliminated in the course of the iterative formative evaluation studies, some technical problems remained unresolved for economic reasons and affected user experience (e.g. problems with missing or mislabeled products in the article data base provided by the home improvement store chain).

The positive findings of the three presented evaluation studies led to the implementation of ten shopping robots in three selected upscale stores of toom Baumarkt GmbH, a German home improvement trade chain. These stores specialize in interior design and are a mix between home improvement store and furnishing house. Therefore, they attract more customers in need of consulting. For this reason, the shopping robots are supposed to handle all the routine questions about article locations, so that staff can focus on providing in-depth interior consulting. This purpose was clearly communicated to both customers and employees to prevent possible concerns of shopping robots replacing humans or leading to job loss among staff.

An economic evaluation of the market-ready shopping robot TOOMAS is still on the agenda. Such a study should aim at clarifying whether and to what degree permanent implementation of robots in home improvement stores is profitable. Production, purchasing and maintenance costs have to be compared to economic profit and added value in a systematical way. Ideally, spontaneous robot use by customers should be documented. Additionally, further examination of user acceptability over a longer period of time promises detailed insights. These do not have to be restricted to customer acceptability. Also, employees' acceptability is of interest: Do they have concerns or fears in relation to the robot? Do they see the robots as competition or surveillance and how can such concerns be reduced? The launch of shopping robots should be accompanied by external and internal strategic communication activities to assist customer and staff acceptability. These PR activities should be systematically planned and evaluated as well.

9 Conclusion

Considering long-term goals like implementation of shopping robots into practice, taking a user-centered approach right from the start of development is essential. Still, comprehensive frameworks for service robot usability testing and robot acceptability need to be developed. However, transferring and adapting approaches from evaluating user interfaces and technology acceptability assessment has proven fruitful in case of our project. The theoretical frameworks used served as a basis to measure usability with subjective and objective measures, as well as taking into account different sub-groups according to sociodemographic variables and computer skills. Thereby, more detailed insights in social and technical aspects of human-robot interaction could be gained. These results formed the basis for creating solutions and ongoing optimizations of the shopping robot TOOMAS and its introduction to an everyday environment like the home improvement store.

Acknowledgments The authors would like to thank Yvonne Ludewig, Gesine Märten, Carina Meixner, Frauke Posselt, Susann Räcke, Jenny Reichenbacher, Uwe Sonderhoff, Julia Wolf, and Cornelia Zinner for their help in organizing and conducting the studies. The authors thank toom Baumarkt GmbH for their support with conducting the field trials and their institutional support with the SERROKON project.

References

1. Bartneck C, Forlizzi J (2004) A design-centred framework for social human-robot interaction. In: The 13th IEEE international workshop on robot and human interactive communication, RO-MAN 2004, pp 591–594
2. Beer JM, Prakash A, Mitzner TL, Rogers WA (2011) Understanding robot acceptance. Tech. Rep. HFA-TR-1103, Georgia Institute of Technology, School of Psychology, Human Factors and Aging Laboratory, Atlanta, GA
3. Bischoff R, Graefe V (2004) Design principles for dependable robotic assistants. Int J Hum Robot 1(1):95–125
4. Böhme HJ, Scheidig A, Wilhelm T, Schröter C, Martin C, König A, Müller S, Gross HM (2006) Progress in the development of an interactive mobile shopping assistant. In: Proceedings of the joint conference on robotics: 37th international symposium on robotics, ISR 2006, and 4th German conference on robotics, Robotik 2006, Paper no. 38
5. Bryman A (2008) Social research methods, 3rd edn. Oxford University Press Inc., New York
6. Butler JT, Agah A (2001) Psychological effects of behavior patterns of a mobile personal robot. Auton Robot 10(2):185–202
7. Datta C, Kapuria A, Vijay R (2011) A pilot study to understand requirements of a shopping mall robot. In: The 6th ACM/IEEE international conference on human–robot interaction, HRI 2011, pp 127–128. doi:10.1145/1957656.1957694
8. Datta C, Vijay R (2010) Neel: An intelligent shopping guide using web data for rich interactions. In: The 5th ACM/IEEE international conference on human–robot interaction, HRI 2010, pp 87–88
9. Decker M, Dillmann R, Dreier T, Fischer M, Gutmann M, Ott I, Spiecker genannt Dömann I (2011) Service robotics: do you know your new companion? Framing an interdisciplinary technology assessment. Poiesis Prax 8:25–44. doi:10.1007/s10202-011-0098-6
10. Gai S, Jung EJ, Yi BJ (2013) Mobile shopping cart application using kinect. In: Proceedings of the 10th international conference on ubiquitous robots and ambient intelligence, URAI 2013, pp 289–291. doi:10.1109/URAI.2013.6677370
11. Garcia-Arroyo M, Marin-Urias LF, Marin-Hernandez A, GdJ Hoyos-Rivera (2012) Design, integration, and test of a shopping assistance robot system. In: The seventh annual ACM/IEEE international conference on human–robot interaction, HRI 2012, pp 135–136. doi:10.1145/2157689.2157722
12. Gharpure C, Kulyukin V (2008) Robot-assisted shopping for the blind: issues in spatial cognition and product selection. Intell Serv Robot 1(3):237–251. doi:10.1007/s11370-008-0020-9
13. Gharpure C, Kulyukin VA, Kutiyanawala A (2006) A robotic shopping assistant for the blind: A pilot study (Technical Report USU-CSATL-1-01-06, Computer Science). Tech. rep., Assistive Technology Laboratory, Department of Computer Science, Utah State University, Logan
14. Göller M, Kerscher T, Zöllner JM, Dillmann R, Devy M, Germa T, Lerasle F (2009) Setup and control architecture for an interactive shopping cart in human all day environments. In: Proceedings of the 14th international conference on advanced robotics, ICAR 2009, pp 1–6
15. Göller M, Steinhardt F, Kerscher T, Dillmann R, Devy M, Germa T, Lerasle F (2010) Sharing of control between an interactive shopping robot and it's user in collaborative tasks. In: The 19th IEEE international symposium on robot and human interactive communication, ROMAN 2010, pp 626–631
16. Gross HM, Böhme HJ (2000) Perses - a vision-based interactive mobile shopping assistant. In: The 2000 IEEE international conference on systems, man, and cybernetics, SCM 2000, vol 1, pp 80–85
17. Gross HM, Böhme HJ, Schröter C, Müller S, König A, Einhorn E, Martin C, Merten M, Bley A (2009) TOOMAS: Interactive shopping guide robots in everyday use - final implementation and experiences from long-term field trials. In: The 22nd IEEE/RSJ international conference on intelligent robots and systems, IROS 2009, p 20052012
18. Gross HM, Böhme HJ, Schröter C, Müller S, König A, Martin C, Merten M, Bley A (2008) Shopbot: progress in developing an interactive mobile shopping assistant for everyday use. In: The 2008 IEEE international conference on systems, man and cybernetics, SCM 2008, p 34713478
19. Gross HM, Koenig A, Schroeter C, Boehme HJ (2003) Omnivision-based probabilistic self-localization for a mobile shopping assistant continued. In: The 16th IEEE/RSJ international conference on intelligent robots and systems, IROS 2003, pp 1505–1511
20. Hornbaek K (2006) Current practice in measuring usability: challenges to usability studies and research. Int J Hum Comput Stud 64(2):79–102. doi:10.1016/j.ijhcs.2005.06.002
21. Ichbiah D (2005) From science fiction to technological revolution. Harry N. Abrams, New York

22. International Federation of Robotics (IFR): Executive summary—world robotics 2013 service robots (2013). http://www.ifr.org/uploads/media/Executive_Summary_WR_2013_01.pdf. Accessed 30 April 2014
23. Iwamura Y, Shiomi M, Kanda T, Ishiguro H, Hagita N (2011) Do elderly people prefer a conversational humanoid as a shopping assistant partner in supermarkets? In: The 6th ACM/IEEE international conference on human–robot interaction, HRI 2011, pp 449–456
24. Jung J, Kanda T, Kim MS (2013) Guidelines for contextual motion design of a humanoid robot. Int J Soc Robot 5(2):153–169. doi:10.1007/s12369-012-0175-6
25. Kamei K, Ikeda T, Kidokoro H, Shiomi M, Utsumi A, Shinozawa K, Miyashita T, Hagita N (2011) Effectiveness of cooperative customer navigation from robots around a retail shop. In: The 3rd IEEE international conference on privacy, security, risk and trust, PASSAT 2011; and 3rd IEEE international conference on social computing, SOCIALCOM 2011, pp 235–241
26. Kamei K, Shinozawa K, Ikeda T, Utsumi A, Miyashita T, Hagita N (2010) Recommendation from robots in a real-world retail shop. In: Proceedings of the international conference on multimodal interfaces and the workshop on machine learning for multimodal interaction, ICMI-MLMI 2010, pp 19:1–19:8
27. Kanda T, Glas DF, Shiomi M, Hagita N (2009) Abstracting people's trajectories for social robots to proactively approach customers. IEEE Trans Robot 25(6):1382–1396
28. Kanda T, Shiomi M, Miyashita Z, Ishiguro H, Hagita N (2009) An affective guide robot in a shopping mall. In: 4th ACM/IEEE international conference on human–robot interaction, HRI 2009, pp 173–180
29. Kanda T, Shiomi M, Miyashita Z, Ishiguro H, Hagita N (2010) A communication robot in a shopping mall. IEEE Trans Robot 26(5):897–913
30. Kidokoro H, Kamei K, Shinozawa K, Miyashita T, Hagita N (2011) You stopped by there? I recommend this:Changing customer behaviors with robots. In: The 13th international conference on ubiquitous computing, UbiComp 2011, pp 569–570
31. Kirchner N, Alempijevic A (2012) A robot centric perspective on hri. J Hum-Robot Interact 1:135–157
32. Kitade T, Satake S, Kanda T, Imai M (2013) Understanding suitable locations for waiting. In: The 8th ACM/IEEE international conference on human–robot interaction, HRI 2013, pp 57–64. doi:10.1109/HRI.2013.6483502
33. Kohtsuka T, Onozato T, Tamura H, Katayama S, Kambayashi Y (2011) Design of a control system for robot shopping carts. In: König A, Dengel A, Hinkelmann K, Kise K, Howlett R, Jain L (eds) Knowledge-based and intelligent information and engineering systems, lecture notes in computer science, vol 6881. Springer, Berlin, pp 280–288
34. Kulyukin VA, Gharpure C (2006) Ergonomics-for-one in a robotic shopping cart for the blind. In: 1st ACM SIGCHI/SIGART conference on human–robot interaction, pp 142–149
35. Kulyukin VA, Gharpure C, Nicholson J, Pavithran S (2004) RFID in robot-assisted indoor navigation for the visually impaired. In: The 17th IEEE/RSJ international conference on intelligent robots and systems, IROS 2004, pp 1979–1984
36. Kulyukin VA, Gharpure C, Pentico C (2007) Robots as interfaces to haptic and locomotor spaces. In: The 2nd ACM/IEEE international conference on human–robot interaction, HRI 2007, pp 325–331
37. Kuo IH, Jayawardena C, Broadbent E, MacDonald B (2011) Multidisciplinary design approach for implementation of interactive services. Int J Soc Robot 3(4):443–456. doi:10.1007/s12369-011-0115-x
38. Leite I, Martinho C, Paiva A (2013) Social robots for long-term interaction: a survey. Int J Soc Robot 5(2):291–308. doi:10.1007/s12369-013-0178-y
39. Liu P, Glas D, Kanda T, Ishiguro H, Hagita N (2013) It's not polite to point - Generating socially-appropriate deictic behaviors towards people. In: The 8th ACM/IEEE international conference on human–robot interaction, HRI 2013, pp 267–274. doi:10.1109/HRI.2013.6483598
40. Lohse M (2007) Nutzerfreundliche Mensch-Roboter-Interaktion. Kriterien für die Gestaltung von Personal Service Robots (in German). VDM Verlag Dr. Müller, Saarbrücken
41. Ludewig Y, Doering N, Exner N (2012) Design and evaluation of the personality trait extraversion of a shopping robot. In: The 21st IEEE international symposium on robot and human interactive communication, RO-MAN 2012, pp 372–379. doi:10.1109/ROMAN.2012.6343781
42. Ludewig Y, Poeschl S, Doering N, Gross HM (2011) Shopping robots in home improvement stores: evaluation of usability and user acceptance of a mobile shopping robot. In: Proceedings of the international HRI pioneers workshop, 2011, pp 66–67
43. Marin-Hernandez A, de Jesus Hoyos-Rivera G, Garcia-Arroyo M, Marin-Urias L (2012) Conception and implementation of a supermarket shopping assistant system. In: Proceedings of the 11th Mexican international conference on artificial intelligence, MICAI 2012, pp 26–31. doi:10.1109/MICAI.2012.21
44. Matsuhira N, Ozaki F, Tokura S, Sonoura T, Tasaki T, Ogawa H, Sano M, Numata A, Hashimoto N, Komoriya K (2010) Development of robotic transportation system - shopping support system collaborating with environmental cameras and mobile robots. In: Proceedings of the 41st international symposium and 6th German conference on robotics, ROBOTIK 2010, pp 1–6
45. Matsumoto T, Satake S, Kanda T, Imai M, Hagita N (2012) Do you remember that shop? - Computational model of spatial memory for shopping companion robots. In: The 7th ACM/IEEE International conference on human–robot interaction, HRI 2012, pp 447–454
46. Mori M, MacDorman K, Kageki N (2012) On the uncanny valley. IEEE Robot Automat Mag 10(2):98–100
47. Mueller S, Schaffernicht A, Scheidig A, Boehme HJ, Gross HM (2007) Are you still following me? Proceedings of the European conference on mobile robots, ECMR 2007, pp 211–216
48. Nakagawa K, Shiomi M, Shinozawa K, Matsumura R, Ishiguro H, Hagita N (2012) Effect of robot's whispering behavior on people's motivation. Int J Soc Robot 1–12. doi:10.1007/s12369-012-0141-3
49. Nakagawa K, Shiomi M, Shinozawa K, Matsumura R, Ishiguro H, Hagita N (2013) Effect of robot's whispering behavior on people's motivation. Int J Soc Robot 5(1):5–16. doi:10.1007/s12369-012-0141-3
50. Nishimura S, Takemura H, Mizoguchi H (2007) Development of attachable modules for robotizing daily items—person following shopping cart robot. The IEEE international conference on robotics and biomimetics, ROBIO 2007, pp 1506–1511
51. Oppermann R, Reiterer H (1997) Software evaluation using the 9241 evaluator. Behav Inf Technol 16(4/5):232–245
52. Pacchierotti E, Christensen HI, Jensfelt P (2005) Human-robot embodied interaction in hallway settings: a pilot user study. In: The IEEE international workshop on robot and human interactive communication, ROMAN 2005, p 164171
53. Pöschl S, Döring N, Böhme HJ, Martin C (2008) Computergestützte Artikelsuche im Baumarkt. Formative Evaluation eines Artikelsuchsystems für mobile Shopping-Roboter (in German). Z für Eval 7(1):113–135
54. Pöschl S, Döring N, Böhme HJ, Martin C (2009) Mensch-Roboter-Interaktion im Baumarkt. Formative Evaluation eines mobilen Shopping-Roboters (in German). Z für Eval 8(1):27–58
55. Pöschl S, Döring N, Gross HM, Bley A, Martin C, Böhme HJ (2011) Roboter-gestützte Artikelsuche im Baumarkt. Eine Studie zu summativer Evaluation und Nutzerakzeptanz eines mobilen Shopping-Assistenten (in German). Z für Eval 10(1):99–132

56. Powers A (2008) Feature: what robotics can learn from HCI. Interactactions 15(2):67–69
57. Reeves B, Nass C (1996) The media equation: how people treat computers, television, and new media like real people and places. Cambridge University Press, New York
58. Schraft RD, Schmierer G (2000) Service robots. Peters, Wellesley
59. Schroeter C, Boehme HJ, Gross HM (2007) Memory-efficient grid maps in RBPF for SLAM using sonar range sensors. In: Proceedings of the 3rd European conference on mobile robots, ECMR 2007, pp 138–143
60. Schroeter C, Gross HM (2008) A sensor-independent approach to RBPF SLAM—map match SLAM applied to visual mapping. In: The IEEE/RSJ 2008 international conference on intelligent robots and systems, IROS 2008, pp 2078–2083
61. Scriven M (1967) The methodology of evaluation. In: Tyler RW, Gagne RM, Scriven M (eds) Perspectives of curriculum evaluation. Rand McNally, Chicago, pp 39–83
62. Shiomi M, Kanda T, Glas D, Satake S, Ishiguro H, Hagita N (2009) Field trial of networked social robots in a shopping mall. In: The IEEE/RSJ 2009 international conference on intelligent robots and systems, IROS 2009, pp 2846–2853
63. Shiomi M, Kanda T, Nohara K, Ishiguro H, Hagita N (2009) Adaptive supervisory control of a communication robot that approaches visitors. In: Asama H, Kurokawa H, Ota J, Sekiyama K (eds) Distributed autonomous robotic systems 8. Springer, Berlin, pp 555–564
64. Shiomi M, Shinozawa K, Nakagawa Y, Miyashita T, Sakamoto T, Terakubo T, Ishiguro H, Hagita N (2013) Recommendation effects of a social robot for advertisement-use context in a shopping mall. Int J Soc Robot 5(2):251–262. doi:10.1007/s12369-013-0180-4
65. Shiomi M, Zanlungo F, Hayashi K, Kanda T (2014) Towards a socially acceptable collision avoidance for a mobile robot navigating among pedestrians using a pedestrian model. Internat J of Soc Robot 1–13. doi:10.1007/s12369-014-0238-y
66. Sucar OI, Aviles S, Miranda-Palma C (2003) From HCI to HRI - usability inspection in multimodal human-robot interactions. In: The 12th IEEE international workshop on robot and human interactive communication, ROMAN 2003, pp 37–41
67. Tokura S, Sonoura T, Tasaki T, Matsuhira N, Sano M, Komoriya K (2009) Robotic transportation system for shopping support services. In: The 18th IEEE international symposium on robot and human interactive communication, RO-MAN 2009, p 320
68. Tomizawa T, Ohba K, Ohya A, Yuta S (2007) Remote food shopping robot system in a supermarket - realization of the shopping task from remote places. In: Proceedings of the international conference on mechatronics and automation, ICMA 2007, pp 1771–1776
69. Tomizawa T, Ohya A, Yuta S (2006) Remote shopping robot system, -development of a hand mechanism for grasping fresh foods in a supermarket. In: The IEEE/RSJ 2006 international conference on intelligent robots and systems, IROS 2006, pp 4953–4958
70. Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: toward a unified view. MIS Q 27(3):425–478
71. Vijay, R., Kapuria, A., Datta, C., Dubey, G., Sharma, C., Taank, G. (2012) Ecosystem of a shopping mall robot—neel. In: Proceedings of the 9th international conference on ubiquitous robots and ambient intelligence, URAI 2012, pp 120–125. doi:10.1109/URAI.2012.6462949
72. Yanco HA, Drury JL, Scholtz J (2004) Beyond usability evaluation: Analysis of human–robot interaction at a major robotics competition. Hum Comput Interact 19(1–2):117–149

Nicola Doering is full professor of Media Psychology and Media Design at Technische Universität Ilmenau. Her research mainly focuses on the social and psychological dimensions of online, mobile and human-robot communication as well as on social scientific research methods and evaluation research.

Sandra Poeschl is assistant professor at the Media Psychology and Media Design Group at Technische Universität Ilmenau. She received her Diploma degree in Psychology in 1998 and her Doctorate degree (Ph.D.) in Communication Science in 2005 at the Ilmenau University. Her research interests lie in psychological aspects of human-machine communication, user experience, user needs, user-centered design and social scientific research methods.

Horst-Michael Gross is full professor for Computer Science with a focus on Cognitive Robotics at Technische Universität Ilmenau and head of the research lab for Cognitive Robotics. He received his Diploma degree (M.Sc.) in Technical and Biomedical Cybernetics in 1985 and his Doctorate degree (Ph.D.) in Neuroinformatics in 1989 at the Ilmenau University. His research mainly focuses on the development of robust and adaptive techniques for mobile navigation and human-robot interaction that cover aspects such as localization, map-building, SLAM, multimodal user detection and tracking, human-aware navigation, situation recognition, dialog adaptation, and several other aspects. As challenging application field, his research is focused on socially assistive mobile robots for public and home environments. He is a member of IEEE, INNS, ENNS, GI and VDI.

Andreas Bley is one of the founders and CEO of the company Metra Labs GmbH. He has studied Industrial Engineering and Management with the focus on automation, intellectual property rights and finance at Technische Universität Ilmenau between 1998 and 2003. He completed his doctoral thesis about the usage of renewable energies by private households at the Institute of Economics, International Economic Relations, of the Friedrich-Alexander-University Erlangen-Nürnberg in May 2007. During his studies he held a scholarship of the Klaus Murmann Fellowship Programme of the Foundation of German Business. Fascinated by interaction of service robots and humans he managed research projects in this topic at MetraLabs, e.g. EU FP7 CompanionAble or EU FP7 Robot-Era.

Christian Martin is one of the founders of the company MetraLabs GmbH, where he is CEO an the Head of Software Development. He has been a Ph.D. student at the Department of Neuroinformatics and Cognitive Robotics at Technische Universität Ilmenau since 2004. He received this doctoral degree in Computer Science and Cognitive Robotics from Technische Universität Ilmenau in 2012. His Ph.D. research was concerned with multi-modal human-robot interaction, especially methods for the detection and estimation of the attention and the interaction interest of a user of a mobile service robot. His interests are Cognitive Robotics (autonomous robots, human-robot interaction, probabilistic robotics) and control architectures for autonomous mobile robots.

Hans-Joachim Boehme received his Diploma Degree in Biomedical Engineering in 1989, his Ph.D. in Neuroinformatics in 1991, and his Habilitation Degree in Computer Science in 2001, all from Technische Universität Ilmenau. Since 2008, he leads the Department of Artificial Intelligence at the University of Applied Sciences Dresden. His research interest focuses on human-robot interaction and applications of service robots in real-life scenarios.