International Journal of Gender, Science and Technology
http://genderandset.open.ac.uk

Navigating Gender Sensitivity in Robot Design: Unveiling the Challenges and Avoiding Pitfalls

Astrid Weiss[1], Sabine Alexandra Zauchner[2], Manuela Ploessnig[3], Nadine Sturm[4], Sofia Kirilova[4], Mathias Schmoigl[3]

[1]Technical University Vienna, [2]MOVES-Center for Gender and Diversity, [3]Salzburg Research, [4]Johanniter Research.

ABSTRACT

Subtle gender cues play a significant role in determining the contexts or tasks to which robots are preferably assigned. This underscores the importance of gender-sensitive design in developing social robots, a key focus of the RoboGen project. In this project, we aimed to design and develop an assistive socially interactive robot that facilitates gender-sensitive human-robot interaction for older adults. While existing research on gender in Human-Robot Interaction is mainly based on experimental studies, where the robot's gender is manipulated and participants' gender assessment is evaluated, we wanted a more contextualized approach. In this article, we report the first phase of the project in which we analyzed requirements for gender sensitivity design through expert interviews as well as a series of workshops with potential senior target users interacting with three different commercially available robotic platforms. Our research demonstrates that gender is a pervasive yet often overlooked factor in Human-Robot Interaction research. Often the gendering process of robots has already taken place before the robots are sold. In our focus groups, we almost found no differences between the three participating groups (women only, men only, mixed) with regard to the two-hour observations of their activities with the robots nor between their statements about the robots. Gender was rarely directly addressed, but circumscribed: It is hidden in statements on the appearance design, voice of robot, name of the robot, tasks the robot is performing etc. Especially with older adults we experienced dominant binary models of gender in their thinking. As also mentioned by our interviewed experts, we see it as our responsibility as researchers to carefully handle these mental models by providing several options of voices (e.g., also a gender-neutral voice), names, tasks the robot performs.

KEYWORDS

gender, diversity, social robots, gender-sensitive design, digitalization

Navigating Gender Sensitivity in Robot Design: Unveiling the Challenges and Avoiding Pitfalls

INTRODUCTION

Gender is a social construct that influences our self-perception, our interactions with others, and our understanding of the world around us, even extending to non-human entities. In the research fields of HCI (Human-Computer Interaction) and STS (Science and Technology Studies), there is a growing awareness of the need to incorporate diversity constructs such as gender (Bardzell, 2010, Robertson, 2017), race (Finda et al., 2020), and their intersections (Schlesinger et al., 2017) in both the process of technology development and research practices. Incorporating diversity constructs becomes particularly crucial when considering humanoid robots, as they are intentionally designed to resemble humans (Lee et al., 2019). Notably, some researchers have highlighted the contradiction that while gender is acknowledged as an important factor in HRI (Human-Robot Interaction), it is often overlooked in the actual design and development phase (Nomura, 2017a; Robertson, 2010; Robertson, 2017; Roff, 2016; Schiebinger, 2019). Failing to acknowledge the impact of gender during the design process risks perpetuating the existing status quo with respect to gender stereotypes (Robertson, 2010) and reinforcing negative aspects, such as those associated with the division of labor (Nomura, 2017b; Robertson, 2017).

In other words, it is crucial to address the issue of gender not only in the appearance design of social robots but also when conceptualizing the interaction. Failing to do so prevents us from envisioning a future path that is sensitive to gender in the important and ever-changing field of companion and care robots. In recent years, several companies launched robotic products, commonly referred to as first-wave social robots, and have impacted our understanding of social companion robots as stationary, speaking entertainment agents. While the promises made by these companies and start-ups were bold, the end products often under-delivered in terms of the actual capacities of the robots (i.e., Weiss, 2021). Among the challenges that these companies have to overcome are technical hurdles and intense competition from other technology-driven products (e.g., speech recognition), which frequently outperform robots in terms of capacity and price, or both. Thus, in this article, we seek to address the following question: What can we learn from the research on first-wave social robots to create more meaningful social companion robots for and with people in the future?

The aim of the RoboGen project is to investigate and implement a gender-conscious approach to robot development with a commercially available platform. Specifically, the project focuses on creating an embodied learning agent that incorporates user feedback to provide gender-sensitive options, thus enabling gender-sensitive human-robot interaction. The primary beneficiaries of this project are senior citizens aged 60 and above who live either at home or in assisted living facilities. The project is divided into four phases: (1) Analysis of gender aspects relevant to human-robot interaction, (2) Development of interaction scenarios, (3) Integration of a learning agent into a commercially available companion robot, and (4)

Evaluation of the robot with the target users. This article provides an overview of the findings obtained during the initial phase of the project – the analysis of gender aspects relevant to HRI. This phase encompassed a thorough literature analysis, three expert interviews, a forum analysis, and requirement workshops involving the target users. Additionally, the article addresses the challenges encountered in translating these findings into the actual design of the robot's appearance and behavior.

BACKGROUND

Socio-demographic issues and new technologies

Higher life expectancy and low birth rates are continually shifting the age structure of our society. The proportion of younger people is falling, while the number of older people is increasing, which changes the shape of the European-27's age pyramid. According to Eurostat (2020a) in 2019, more than one fifth of the EU population was aged 65 and older and they expect the share of people aged 80 years or more to more than double by 2100 covering 14.6 % of the whole population. This is likely to be of major significance in the coming decades not only with regard to the financing of pensions, but also concerning care and nursing of elderly people. Though the increase in life expectancy is evident for both genders, women in Europe in 2018 reached an average age of 83.7 years while men reached 78.2 years - a mean gap of 5.5 years (Eurostat, 2020b). Consequently, about 2/3 of 80 to 84-year-olds and about 3/4 of people aged over 84 are women (Eurostat, 2020b).

This means that women at the age of 65 in Europe can expect to live a smaller proportion of their remaining lifespan in good health without functional limitations or disabilities, than men of a similar age. While women spend 47% of their remaining lifetime in health, this figure is 54.1 % for men. The shortage of skilled healthcare workers has been a massive problem for societies for years, exacerbated recently by the Covid-19 pandemic. The skills shortage is expected to continue unless structural deficits in this female-dominated working field are addressed including low pay, weak career development, poor lifelong learning, missing social recognition, and the neglect of emerging technological innovation (Michel & Ecarnot, 2020; Weber, 2017).

So how might technological change help with this problem? Autonomous "things" are number eight of the Gartner Top 10 Strategic Trends for 2020 (Cearley, 2020) using Artificial Intelligence to automize functions previously done by humans. The rapid developments in this field make the use of social robots seem plausible as a means of counteracting the lack of qualified staff. Socially interactive robots, according to Dautenhahn (2007) are diversely defined, depending on the particular focus of research for which the social functions are developed:

> "(…) express and/or perceive emotions; communicate with high-level dialogue; learn models of or recognize other agents; establish and/or maintain social relationships; use natural cues (gaze, gestures, etc.); exhibit distinctive personality and character; and may learn and/or develop social competencies." (p. 684)

When it comes to the design and development of social robots, there is no reason to believe that the gender imbalance among IT specialists – defined as those employed across all industries who can develop, operate and maintain IT systems – would be any different in robotics than in general: In total, 17.9% of IT specialists in the European Union are female and 82.1% male (EIGE European Institute for Gender-Equality, 2018). This gender imbalance has implications for the design of new systems and artifacts. Gender research in technology has been aware for decades that technological artifacts are not neutral, but they transport social ideas, attitudes and values, which in turn manifest themselves in the products or are inscribed in them (Akrich, 1992).

Gender stereotypes

In the approaches of (de)construction (Butler, 1991, 2004; Villa, 2010), the focus of interest is on the production of social gender, on "doing gender" (West & Zimmermann, 1987; Wetterer, 2010). The term "gender" here refers to everything that is considered typical in a culture and at the same time untypical for a particular gender. This includes the construction of gendered attributions of certain behaviors, interests, competencies, or attitudes, and the hierarchizations and social power aspects associated with them. Gender is seen as an actively shapable and basically also context-related, situation-specific and de-constructable context of action (Gildemeister, 2010; Gildemeister & Wetterer, 1992).

Taking into account the concept of the co-construction of gender and technology established in gender research (Wajcman, 2000, 2004, 2010), in which it is assumed that gender and technology are in a reciprocal, flexible, and malleable relationship to one another, gender stereotypes must be handled with particular sensitivity in the construction of social robots, so that these stereotypes are not arbitrarily transferred to the new technology (see also Akrich, 1992). As Tatsuya Nomura states in his summary of studies on gender and robotic technology, "gender characteristics are one of the most important considerations in robotics design to influence interaction between robots and humans." (Nomura 2017a, p.10). That gender aspects not only influence human-to-human communication but have a similar effect in human-robot interactions has been shown in several studies. One example is the study by Eyssel & Hegel (2012) which examines whether subtle gender cues in anthropomorphic robots influence users' perception of the robots and the tasks they are performing. The study of 60 students (30 women/30 men) shows that robots are perceived differently depending on the length of their hair: The robot equipped with chin-length hair was rated as female and the robot with short hair as male. These assessments were associated with stereotypical characteristics (e.g., determinant vs. social) and suitability for certain activities (e.g., domestic support vs. repair of technical equipment), which were attributed to the sexes. "(...) the robots' hair cues (apparently activated), participants' knowledge structures about males and females, and gender stereotypes subsequently biased the evaluations of the robots." (p. 2223).

Tay et al. (2014) came to similar conclusions for a student sample (84 men; 79 women; 1 undisclosed) aged 20-35 years, when they found that "male" robots were found to be more suitable for a safety scenario and "female" robots more suitable for a health scenario. They further note that participants in the study found gender-specific social robots more easily accepted if they conformed to their respective occupational role stereotypes: "Participants more easily accepted social robots with gender and personality that conformed to their respective occupational role stereotypes." (Tay et al., 2014, p. 82). Similarly, Stroessner & Benitez (2019) came to the conclusion in two recent studies (Study 1: 145 men and 104 women aged 18 to 73 years; Study 2: 106 men, 85 women and 5 undefined, aged between 19 and 71 years) that robots were rated more positively and generated a greater desire for contact if they were considered human-like and female. Another study by Carpenter et al. (2009) revealed that – according to stereotypes – the study participants prefer female robots for use in the home.

Thus, gender aspects not only influence human-to-human communication but also have a similar effect in human-robot interactions. Taken together, this would speak strongly in favour of using gender-neutral robots, as recommended for example by Nomura (2017b) or Schiebinger (2019). However, as we will reflect later in relation to our study findings, we should consider whether a social robot will ever be perceived as gender-neutral.

Interaction and Contexts

When it comes to the interaction effects of gender with other social categories such as ethnicity, theories of intersectionality (Crenshaw, 1989; Lenz, 2010; Walgenbach, 2012, 2013) offer themselves as explanatory models. Intersectionality assumes that social categories cannot be viewed in isolation from one another but must be analyzed interwoven or in their "intersections". Stereotypical notions of men and women thus interact with equally largely unquestioned notions of, for example, people with a migration background, different ages, low education, lower social status or sexual orientation (Lenz, 2010). Therefore, characteristics other than the gender of the robot also influence the evaluations and perceptions of them by the target groups. Personal characteristics of people and contextual or situational factors (e.g., robots at home, in nursing homes, in the professional environment) also play a significant role in HRI.

Strait et al. (2015) expanded the target group in their study in comparison to the above-mentioned studies, both by the size (510 persons, 62% of them male) and by the age categories included (18 to 75 years) and examine the influence of a politely praising language of the robot when instructing for a task. Robots using a polite speaking style were generally perceived more positively. There were no age-related differences when it came to the evaluation of polite robots, although women valued them more than men. Mutlu et al. (2006) also found that men - again in accordance with gender stereotypes - achieve better results when they compete with robots in a computer game than when they work collaboratively with the robot. Women, however, achieve the same results in both experimental conditions.

A similar study on the interaction of robot and participant gender and activities was conducted by Kuchenbrandt et al. (2012). In an experiment on 35 female and 38 male participants in Germany, the effects on the performance of men and women in a sorting task on a touch screen were investigated depending on whether the instructions were given by a male or a female robot. The results showed that the female participants completed the task at the same speed regardless of the gender of the robot, but that the male participants completed the task faster when instructed by a “male” robot.

Siegel et al. (2009), in turn, found a cross-gender effect in a science museum in Boston (USA) with 76 men and 58 women. There were three robot versions - neutral, female and male, each defined by language. After a brief interaction with the robot and an appeal for donations, the results showed that men were more likely to donate money when asked to do so by a female robot. Among the men, the trust and assessment of commitment were also significantly stronger when asked to donate by a female robot. Women showed little preference. However, the participants tended to rate the robot of the opposite sex as credible, trustworthy, and committed. These results show how important it is to consider both robot and human gender when designing for human-robot interaction.

In the study by Alexander et al. (2014) the participants (24 women and 24 men, the age is not specified - the "Yale community" is spoken of) had the task of solving Sudoku-like puzzles with the help of a male or female robot (defined by voice and name). Contrary to expectations, participants asked the robots for help more often, regardless of their gender. Participants of both sexes, in turn, experienced the above-mentioned “cross-gender effect” and found that they felt more comfortable with a robot of the opposite sex and that they preferred the help of the male robot.

Coming back to gender neutral design of social robots, an interesting approach is offered by Dufour & Ehrwein Nihan (2016). In their article they ask the question whether robots need to be gender-specific and thus stereotyped in order to increase the acceptance and economic value of the machine. Instead of giving the gender-specific look or voice as an indication of how the robot is perceived, they highlighted its technical characteristics when asking for users' evaluation. Although the authors assume that there is a need for further research, their preliminary result shows that "the effect of human stereotypes on the judgments of robots is not inevitable" (ibid., p. 8), since the evaluation or categorization of the robot was performed by the participants along the given technical characteristics.

Despite some fixed cornerstones, research on gender aspects of social robots is mainly based on experimental studies, where the robot’s gender is manipulated, and participants’ gender assessment is evaluated. Often these studies consider the “human in the vacuum” and are lacking contextualization (Lee et al. 2022). It also shows that - as so often in early research - the focus of studies on gender and robotics is mainly focusing on the differences in a binary understanding.

This is why it is so important to employ a gender-sensitive approach that equally takes the needs of all genders in the design of a social robot solution into account, especially when critical feminist research in this field is just beginning to establish itself. We need to find ways to incorporate gender research without reducing it to test for perception differences.

RESEARCH APPROACH

We based our gender-conscious approach for the first phase of the RoboGen project on the CASA (Computers-Are-Social-Actors) paradigm proposed by Lee and Nass in 2010. This paradigm assumes that the social strategies employed in human interactions, along with their associated gender stereotypes, are also applied to interactions with computers and robots. Our primary focus was on investigating the following research questions:

1. What factors should be taken into account regarding gender sensitivity in social robot design from an expert's perspective?
2. What requirements and preferences do seniors express concerning the gender of a social robot?
3. How can these requirements be transferred into actual appearance and behavior/interaction design?

However, addressing gender sensitivity encompasses more than just expert and target user involvement, it requires the entire team to be capable of acting in this regard. Therefore, the initial step in our project involved developing an understanding of gender that is oriented toward deconstructive and intersectional models. Additionally, the interdisciplinary project team continuously engaged in reflection on gender stereotypes, fostering an environment conducive to ongoing dialogue and critical analysis.

We started our project by exploring a range of commercially available robot embodiments through an online forum analysis. Next, we conducted expert interviews in order to inform ourselves about the current research topics related to our project from different scientific perspectives. Recognizing the significance of involving the primary target group in the design process, we conducted interactive focus groups with prospective users, as observed in studies by Ahmadi et al. (2018), Brenssell and Lutz-Kluge (2020), Daugherty and Wilson (2018), and Hartung et al. (2020). Subsequently, the research team collaboratively created interaction scenarios. Building upon the importance of placing technology users at the center of attention, we align with the viewpoint expressed by Messmer and Schmitz (2004), stating that our approach to gender and ICT aims to adapt technology to user-oriented demands rather than forcing users to conform to the technology (p. 248).

Robot Embodiment Analysis

In the RoboGen project, our focus was on the segment of commercially available social robots that fall within the price range of approximately €200 to €400. This price range was chosen considering the financial accessibility for senior citizens in the relevant contexts (home and care facility). Below, we provide a brief description of the three selected embodied digital assistants that were chosen to be modified for the project's objectives.

Figure 1: Q.bo One

Q.bo One (see Fig. 1) is an “Open-Source Personal Robot Assistant” developed by the Spanish provider TheCorpora. This stationary humanoid robot is designed to enable the introduction of assistance robots into everyday life. It is easily programmable and can be connected to language assistants such as Siri or Alexa, offering numerous application possibilities for technology-affine end users at an affordable price.

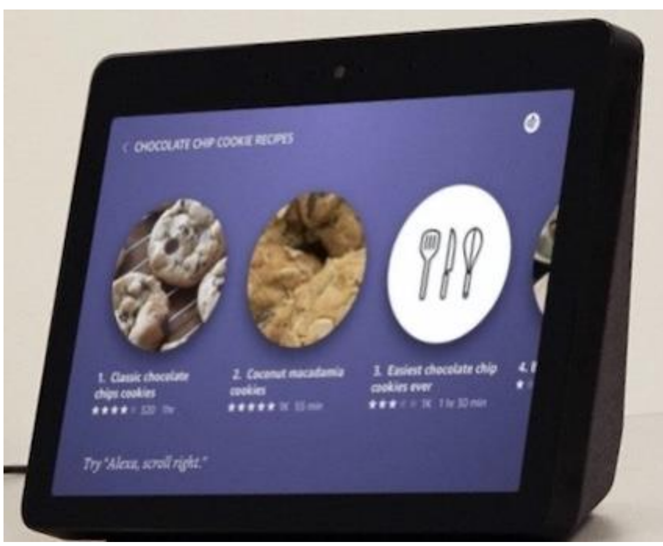

Figure 2: Echo Show

Echo Show (see Fig. 2) is a virtual assistant display. It is an Amazon product, a 7-inch touchscreen with speaker, equipped with Alexa functionalities. Echo Show supports audio as well as touchscreen input via a display. Echo Show can be combined with Q.bo One and with Anki Vector.

Figure 3: Anki Vector

The Vector robot, developed by Anki in 2018, was intended as a social companion for homes (see Fig. 3). It is designed to be both a helpful assistant and a friendly companion for people at home. Equipped with speech recognition and synthesis capabilities, Vector can answer questions and engage in conversations. Its expressive LCD eyes and use of a single-point laser for map creation enable it to navigate its surroundings using SLAM (Simultaneous Localization and Mapping) technology. Additionally, Vector can be integrated with Alexa for added functionality. Onboard, a convolutional neural network facilitates people detection and other tasks. Compared to the stationary head-and-torso design of Q.Bo One, Vector's design aims for a more creature-like and dynamic experience. Constantly moving around and emitting sounds, it creates a life-like presence within the household.

As part of our analysis phase, we conducted a comparative content analysis in online forums to examine the use of anthropomorphic language for all three systems. We also explored how gender perceptions might influence these anthropomorphic impressions. Initially, we expected to find the highest degree of anthropomorphic language use for Q.bo One due to its humanoid shape; however, "findings suggest that the life likeness of the artifact is not pre-dominantly linked to the appearance, but to its interactivity and attributed agency and gender" (Weiss et al., 2020), so it was Vector that was talked about with the highest degrees of anthropomorphic language. This preliminary investigation already provides valuable insights into the complexity of anthropomorphism in connection to different materializations/ embodiments of voice assistants. There is a high impact of apparent agency and voice design, rather than the physical embodiment we initially expected to be influential.

Related research has provided similar findings, including a study that examined the perceptions of university students toward the voice assistant Alexa. Through an online survey, participants regarded Alexa as a distinct entity, attributing both technical and social characteristics to the system (Fortunati, Edwards, et al., 2022).

Similarly, a YouTube video analysis of the Sophia robot showed that its highly human-like design did not significantly influence the participants' perception of the robot as human-like, however the use of Sophia was found to be associated with gender stereotypes typically associated with women's professions and occupations, although not entirely (Fortunati, Manganelli, et al., 2022). For more in-depth information on our forum analysis, please refer to Weiss et al., (2020).

Expert Interviews
Following our initial analysis of online forums, we recognized the importance of further exploring gender-conscious social robot design through an expert lens. To gain comprehensive insights, we conducted expert interviews in August and September 2019. We had the privilege of engaging with renowned international experts, each with their unique expertise and perspectives, who are listed below in the order that the interviews were conducted:

- Ex1. Londa Schiebinger, Professor of History and Director of the Gendered Innovations Project of Stanford University, CA USA, is known worldwide in the community. Professor Schiebinger approaches the topic from a gender perspective.
- Ex2. Tatsuya Nomura, Professor at the Department of Media Informatics at Ryukoku University in Japan. Professor Nomura contributes his expertise primarily from the perspective of human-machine interaction and robotics.
- Ex3. Astrid Rosenthal-von der Pütten is a Professor at the Department of Society, Technology, and Human Factors at the Human Technology Center and RWTH Aachen University. She, in turn, looks at the topic mainly from a media-psychological perspective.

The diversity of disciplinary backgrounds among these experts proved to be particularly advantageous. Not only did they bring their gender expertise, but each contributed complementary specialist skills, adding depth and richness to our exploration of gender-conscious social robot design.

Interactive Focus Groups
To understand the specific requirements and preferences that seniors have concerning the gender of a social robot, we opted to conduct interactive focus groups. We were convinced that engaging in interactions with actual systems would provide us with more authentic insights, capturing real-world concerns and genuine desires, rather than relying on hypothetical or over-exaggerated expectations (Weiss & Spiel, 2022). By involving seniors in these interactive sessions, we aimed to gain practical and valuable feedback to inform the design and development of gender-sensitive social robots that truly cater to their needs. The interactive focus groups were held in November 2019 at a seniors' residence in Vienna, Austria. The session was divided into two parts. Firstly, the participants familiarized themselves with and explored the three robot embodiments, with each system having a dedicated "interaction station" featuring a total of 33 simple use cases participants could explore. Each station had the support of at least one project team member, and two team members observed the activities of the participants. This phase lasted for two hours, followed by a one-hour focus group discussion.

During the exploration phase, the participants were given the freedom to choose which use cases they wanted to try out, with each station having a specific focus:

Echo Show: The use cases involved examples of Alexa skills, daily routines consisting of two Alexa skills, and structured dialogues related to sleep, stress, sports, or games.
Q.bo One: The use cases demonstrated that the robot can recognize feelings from spoken text and from facial expressions, and participants could explore different voices.
Anki Vector: The use cases included examples of Alexa Skills and structured dialogues related to sleep, stress, sports, or games. Notably, the Vector commands were still in the English language (e.g., time).

Three focus groups were conducted in total, with the first having five participants and the second and third each having six participants. One group was comprised exclusively of women, one of men, and the third group was mixed, consisting of three women and three men. Before the study, the participants received a questionnaire to provide information on various factors such as demographics, educational level, health status, family status, and their affinity for and regular technology usage.

In the **women's focus group**, the participants' ages ranged from 63 to 87 years, and all were retired. The level of completed education varied from compulsory schooling to university degrees. Three participants reported having chronic diseases, and three were widowed, while the others were either in a partnership or divorced. Classical media such as television, radio, and reading were used frequently, along with regular private use of the Internet. Two participants played online games several times a day or daily. Mobile phones were the primary mode of telephony, and the alarm clock on mobile phones was frequently used. Two participants used telemedicine several times a day. All participants had Internet access at home, and two of them used intelligent thermostats for heating.

In the **men's focus group**, the participants' ages ranged from 69 to 77 years, and all were retired. The highest education level varied significantly. Two participants reported suffering from chronic diseases. All participants were either married or in a partnership. Similar to the women's group, there was frequent use of classical media along with PC/notebook and computer games. Mobile phones were more commonly used for telephony than the landline network, with functionalities such as photography, calendar, and alarm clock being frequently used. All participants reported using telemedicine services several times a day and had an Internet connection at home. Some participants used a thermostat for household automation, and one participant had a lawn-mowing robot.

The **mixed focus group** had participants aged from 63 to 87 years, with diverse educational backgrounds, and all were retired. Two participants reported suffering from chronic diseases. Family statuses in this group were diverse. Media usage showed frequent use of traditional media, along with computers (PC, notebook), and moderate use of online shopping was indicated. Most participants used mobile phones for telephony, but video telephony was rarely used. Five out of six participants used telemedicine services several times a day. All participants had Internet access at home, and three of them used home automation, such as thermostats, room brightness control, and home robots.

Analyses

The expert interviews and the interactive focus groups were analyzed according to the qualitative content analysis of Mayring (1990, 2002), which offers a structured methodology for the analysis of qualitative data (also the online forum study was analyzed following this methodology). This model follows the process of (1) transcriptions of the audio recorded data (2) paraphrasing of the content regarded to be relevant for the research questions (3) creating a system of categories (4) continuous reviewing of the category system, (5) summarizing and structuring, (6) interpreting and (7) back-checking of the source material. Throughout the analysis, our primary focus was on identifying similarities and differences in the way thematic priorities were addressed, shedding light on the various perspectives and viewpoints expressed by the participants. This rigorous methodology allowed us to derive valuable conclusions and implications from the data, providing a deeper understanding of the experts' as well as the seniors' perceptions and attitudes regarding gender-sensitive social robot design.

FINDINGS

Expert Interviews

The following section provides concise summaries of experts' insights regarding the essential factors they deem crucial for gender-sensitive design in social robots. It is important to note that these summaries are intentionally concise (excluding direct quotes) to maintain the paper's overall length.

During the interviews, the question arose about *whether it is desirable to have different communication approaches for men and women in social robots*, considering the potential for perpetuating gender stereotypes. Ex3 highlighted two key aspects: the design of robot interaction behaviour to cater differently to men and women, and questioned whether such gender-specific strategies are necessary or wanted. Ex2 responded by suggesting that gender stereotypes can be avoided by focusing on the specific target user group. The better we understand their requirements and needs, the more evident gender sensitivity becomes. Ex2 further emphasized the importance of knowing the gender bias that exists among the targeted people, such as their age, education, and medical conditions. Gender sensitivity for virtual assistants was considered a must-have. Ex2 pointed out that failure to consider gender aspects in robot design could lead to rejection by users. Ex2 also acknowledged differences within the target group, particularly related to physical conditions, but did not emphasize gender as a differentiating factor. Ex3 recognized differences among target groups but did not refer to gender dichotomies and highlighted the need to classify users based on factors like flexibility, enterprising nature, and attachment to self, indicating an intersectional understanding of gender.

Regarding *methodological challenges*, Ex2 stressed the complex interaction effects of gender with other factors like robot gender, age, and context, underscoring the need for caution in handling these complexities. Ex3 found it difficult to take a gender-sensitive approach in development due to existing biases. Foreseeing potential gender biases, such as male-dominated engineering teams influencing face recognition technology, presented challenges. Consistent with these perspectives, Ex2 recommended the use of gender-neutral robots, considering them essential for gender sensitivity and acknowledged that some researchers may develop gendered robots, but emphasized the significance of gender-sensitive designs. Further, the expert emphasized the relevance of gender sensitivity from an ethical standpoint, recognizing the need for inclusivity in targeting people. Ex1 highlighted *the importance of using gender-neutral names for robots* to avoid perpetuating stereotypes, and emphasized the need for gender-inclusive language, treating both genders with equal respect and providing users with a wide range of voice choices.

It can be concluded from the above statements that the experts mainly propose a deconstructive (Butler, 1991, 2004; Gildemeister & Wetterer, 1992; Wetterer, 2010) understanding of the genders, in one case intersectionality (Crenshaw, 1989; Walgenbach, 2012, 2013) can also be assumed. However, we also observed statements that still indicated tendencies towards "difference approaches" that focus on binary gender notions.

Focus Groups

In the forthcoming paragraphs, we will outline the outcomes derived from the focus groups centered around the Q.Bo One and Vector robot. These results will be categorized into two aspects: (1) behavioural observations during participants' familiarization with the systems and (2) an overarching evaluation of the systems based on the collective insights emerging from the focus group discussions.

Getting to know and testing the systems – Observations

Volume and intelligibility: The preset volume of all three systems was too low for the target group of senior citizens. Only at the Echo Show could the volume be turned up. However, this had the side effect that Echo Show could no longer understand the wake-up word for some skills. The voices of Q.bo One and Anki Vector could not be turned up, which made the spoken texts partly incomprehensible. Vector was especially difficult to understand even in its English source language, also because of its fast way of speaking.

Addressing the systems: In certain instances, both systems showed a lack of response to the designated wake-up word for various reasons. On occasion, the word was uttered too softly, while in other cases, Vector was distracted by the presence of faces. Furthermore, participants frequently encountered challenges in maintaining the pause following the wake-up word. This deviance from how we usually talk requires adjusting when using these systems.

Decision trees: These provide users with potential reasons and solution choices for a given problem (e.g., What can I do when I feel lonely). We observed that listing several options on the Echo Show, which communicates with the participants linguistically and via text on the monitor, was a usable way for them to interact with the system. However, in the case of Vector, which exclusively delivers text in verbal form, participants tended to lose track and felt overwhelmed by how to continue the interaction.

Recognition of emotions: (1) Facial expression recognition: This task exhibited a notably low success rate. We assume that employing an enhanced camera (as supported by preliminary tests on other devices) could potentially yield superior outcomes. Suboptimal lighting conditions further contributed to the diminished accuracy of the results. Nevertheless, a considerable number of participants derived a substantial sense of enjoyment from engaging in this activity. (2) Text-based recognition: The way Q.bo One was activated diverged from the conventional wake-up word approach. We used a dedicated app for control, making Q.bo One quickly responsive and attentive. At first, this approach caused confusion among our participants, as it deviated from the handling employed by Echo Show. However, after a short demonstration with a sample sentence, the task was easy to master for the participants. An intriguing point raised by one participant was the potential for Q.bo One to identify panic (a form of fear) as an emotion and then initiate help. Furthermore, in cases where pain is detected, the participant suggested notifying a designated contact person if deemed necessary.

Voices: Four distinct voices were provided for Q.bo One and the participants were requested to provide their feedback. The evaluations varied significantly both among different groups and within individual groups. However, overall, seniors specifically emphasized the importance of choosing a voice that is easily understandable as a crucial factor for communication and addressing gender-sensitivity as secondary.

Recommendations for future development: Regarding the Echo Show there was notable interest in the option to input emergency contacts that the system could contact in case of an emergency. Additionally, regarding Q.bo One, there was a desire expressed for the robot to have the capability to contact a designated reference person when needed.

Evaluating the systems – Focus Groups

Several of the above-mentioned aspects were reiterated during the follow-up discussions with the seniors. As anticipated, there was a consensus that the systems are not yet ready for the market, particularly when considering our target audience and the presented versions of the task implementations. In the words of FG2_M5, “For me these are building blocks, research building blocks to a future larger whole.” FG2_M4 expressed, “Wasn't very practical, it didn't work much, that's all future, still has to be developed.”

The prevailing sentiment among participants was that the devices are geared toward individuals with limited mobility. Therefore, any potential purchase would largely depend on the life situations of the users. As FG2_M13 stated, "If I was badly on foot and could not drive around as I do now, I would certainly take something like this. Because that is where I have speech, that is where I can do something." In contrast, FG3_F1 shared, "This is for very sick people who are in a wheelchair, but as long as I can still move and stir, thank God, I'll look it up myself or look in the newspaper. (...) Otherwise I won't get up at all."

Consequently, the present value added by the systems is still perceived as relatively modest. As FG2_M11 questioned, "The question is, what can Google do that it cannot do?" FG1_F4 opined, "This system exists, and if you do that, it will have to be a special application for something that does not yet exist. So just duplicating something and casting it into a different garment, that is not sensible, I think." FG2_M15 added, "Yes, for example, the Anki Vector can pick up the cube - it's a funny thing, but what added value should we have?"

Notably, participants shared considerations beyond hedonic novelty, such as practical utility enhancements. FG3_M3 highlighted, "If I can say: 'Find me a phone number' and he gives me the phone number and I print it and can phone. I don't have to search long and hard in the telephone book." FG1_F1 suggested (after longer consideration), "With the robot itself, it would be interesting (incomprehensible) that it helps in some way - when it recognizes you are sad that it can do something, it cheers you up." Finally, FG2_M21 added, "Now to the health parameters: a system that is supposed to be helpful has to collect all parameters that I can test now and make them available for a possible call." Additional ideas emerged during focus group 2 discussions, including suggestions such as remotely activating appliances like cookers and washing machines, and performing tasks like sweeping and handling laundry. The discussion groups also emphasized, in alignment with the description in section 5.3.1, the potential benefits of aiding in emergency situations.

In terms of usability, the speech volume of the devices, particularly Echo Show, was a major point of concern, as noted by FG1_F8, "What bothers me is that it's incredibly quiet, you can't turn it up either, it's the highest volume, and in my opinion it's also very quiet." Additionally, the flexibility of language options and font sizes was deemed essential for user-friendliness, as FG1_F17 mentioned, "You should be able to choose at least 2 languages... and the font, the system must have a sufficient font size, because if I need reading glasses, it is difficult." Moreover, voice control and accent recognition emerged as areas in need of improvement, as participants noted that systems should be capable of understanding a broader range of words and accents (FG3 M8, "It should also recognize the different accents, for example of street names in dialect").

However, despite usability concerns, the women from focus group 1 felt capable of operating the devices independently. This perspective was not universally shared by the participants of the other groups and is contingent upon the acknowledgment that they lack insight into the extent of preparation or configuration required for the demonstrated systems. As stated by FG2_M2, "No, I do not have that feeling." FG2_M3 added, "I don't know if it works with speech only or if it was switched on before or in some other way because we got it ready for use." The Echo Show was perceived as most likely intuitive, assuming the correct wake-up words were utilized. Nonetheless, participant FG2_M6 underscored the devices' lack of alignment with seniors' needs: "Everything that has been shown so far has nothing to do with seniors - so no help for seniors at all. It is - how shall I put it - generally high-tech".

A significant portion of the tasks performed received an average rating ranging from very good to good. The reasons for the comparatively lower ratings of certain tasks, particularly within the first focus group, are elucidated by participants' explanations: FG1_F10 noted, "Well, when I watch the news now, I go to the computer, I don't need that." FG1_F12 mentioned, "The offered food is not ready for use." FG1_F2 expressed, "If I get 10 things suggested now, I will be struck by the information, I don't know what I wanted anymore. Like it was with the recipes or with the restaurants." FG1_F11 pointed out, "So I know myself if I am grumpy or not". Fundamental doubts about the efficacy of emotion recognition emerged as well. As FG2_M10 highlighted, "Everybody expresses pain differently, when I have pain, I withdraw and another one moans all day long - there are different forms of expression. It is difficult to capture this on a picture screen." Even in the context of stress management, the complex decision trees employed by Anki Vector did not resonate positively, largely due to the need for active engagement. In fact, the fundamental utility of decision trees was questioned: FG2_M21 asserted, "I would never offer information trees, because after half a year this information is uninteresting." Furthermore, the daily joke feature from Echo Show received a lukewarm reception. When asked about their enjoyment of working with the robots, participants expressed a mix of sentiments. FG3_F9 reflected, "I think it is something new for our age group, it is impressively natural, I also think it takes getting used to... curiosity is more in the foreground." FG3_M6 responded, "Well look, then we will see. We'll have a look at it now, but it's funny, it didn't happen to me." Despite the varied reactions, participants generally found the interactions enjoyable. As FG3_all summed it up, "You just talk, you watch, and it doesn't work, that was rather fun."

When asked about whether participants perceived that they had gained new insights or learned new skills, the responses exhibited a range of viewpoints, spanning from: FG3_F26's perspective, "Realizing what already exists, what one does not know yet." To FG3_M11's observation, "I have learned that the things that are being discussed and have been discussed for years have reached us and that we can still expect a lot.", to the contrasting view of FG3_M3: "There was nothing here in this room that I had not already heard or seen; actually, that is known." The experience prompted one participant to develop a heightened affinity for technology: FG2_M23 expressed, "I have already realized that with my counterarguments, I am now more open to the idea that when I am in a situation where I need it, I have already become more willing to embrace technology." The discussion also delved into the diversity among individuals: FG3_M10 acknowledged, “Yes, people are very different; everyone desires something different” (met with laughter from the group).

Regarding data protection, the discourse among participants revolved around opposing viewpoints: “I might find it useful under specific circumstances, like limited mobility or loneliness,” and “I wouldn't purchase something like this because I'm uncertain about how my data is handled.” As FG2_M9 shared, “For me personally, it would also be important to have a switch to turn it off, so I know nobody is listening to me now.” However, FG2_M10 offered an alternative viewpoint: “But it should not be switched off for caring relatives.” FG2_M16 emphasized the need for a safeguard: “You would have to ensure that the information remains within an internal circle and doesn't get shared with Alexa.” Contrasting opinions were also voiced, as articulated by FG2_M19: “I just think that the people who need it are happy that they... I don't care who is listening.”

DISCUSSION

With the RoboGen project we aimed at getting into action with gender-sensitive design and development of social robots. The analysis phase of our project, which is presented in this article gave us insights on critical aspects with respect to gender stereotypes which can be reinforced through design from the perspective of experts as well as potential end users.

Gender sensitivity

The gender-theoretical approaches implicitly mentioned by the experts ranged from the perspectives of focusing on differences to (de)construction and intersectionality. Basically, all three experts were against re-stereotyping of robots. This is by itself not an easy process, when considering that gendering processes already have taken place before the robots are sold.

As we also revealed in our forum analysis about the three systems that were used (Weiss et al., 2020), anthropomorphizing does not only take place through gender attributions, but already through the appearance design of the robot, its marketing, and its technology readiness. This was particularly evident in the Anki Vektor forum when the robot was regarded as an animal worthy of protection or as a male family member that triggers parental functions in the users. Otherwise, Q.bo One was mainly discussed as a device/object and continuously addressed as "it". This suggests that the less a virtual assistant is technologically mature the less gender is attributed, which is supported by the study of (Dufour & Ehrwein Nihan, 2016) who assume that a focus on the technical description of a robot helps to avoid gender stereotypes.

Further, the importance of a sensitive handling of gender stereotypes regardless of the status of development is shown by the manufacturers' posting, which clearly gives Q.bo One and Anki Vector a male gender.

Thus, if stereotyping takes place unreflectively at such an early stage, it might help to avoid any subtle gender cues such as the hair length as tested by Eyssel & Hegel (2012), or – as recommended by one expert – giving the robot a gender-neutral name like Chris which in German speaking countries can be a short for Christine (f) or Christian (m). However, this can only be part of the solution. An essential aspect needs to be considered namely the development of a common understanding of what is understood by the term gender within the development team. This would enable a reflexive process to evolve in the course of developments, in which single decisions are consistently questioned against the background of this understanding of gender.

Even if it may seem temptingly simple to treat this via dichotomies this will not be beneficial for technology adoption, as gender-sensitive design needs to reflect, address, and value the diversity of the user target group to create meaningful technology (Lee et al., 2016; Neven, 2010). If this is not the case, robots will most likely not be integrated into everyday life. The attributions, for example, that men would use rather aggressive/commanding language and women would use rather polished/bidding language, would be one of these pitfalls. In fact, such a categorization would include people who act in traditionally gender attributed ways, but exclude people who identify as non-binary, or wish to express themselves politely as men or demandingly as women.

Considering differences

The experts interviewed have repeatedly pointed out that the aim must be to recognize the diverse needs of individuals in their specific contexts, and not of some imagined groups, which would further exacerbate dichotomies.
“If one tries to describe the diversity and differences of individuals with certain characteristics, there is a danger of reducing the complexity of the diversity that makes a person unique”. (Abdul- Hussain & Baig, 2009, p. 27). This means that focusing on such categories without considering the differences that also exist within the categories leads to stereotypes, as Stuber and Achenbach (2004, p. 18) explain: "A specific risk arises when the selection of criteria is reduced and when at the same time the difference in the sense of dichotomy and separation is emphasized." It is precisely in these first development steps that it is important to define exactly such criteria in order to use them in technological development and that the reflection on potential dichotomies must be taken very seriously. It is, therefore, a matter of including other diversity factors, such as the age of users, their physical health, their mobility behavior, their respective needs for support, their service requirements, or their need for entertainment.

Another essential statement of the experts was that the acceptance of users does not depend so much on their technical competence, but that it improves with the concrete use of a virtual assistant. Nevertheless, technical competence and educational background are still regarded as important criteria because these factors are essential when exclusion criteria are spoken of even before a direct contact with a robot.

This became evident in the focus groups. We did not find any differences between the three participating groups (women only, men only, mixed) with regard to the two hours observations of their activities with the robots, nor did we find differences between the statements of the three groups in the focus groups. There were just two exceptions which are far from implying a gender-theoretical concept of difference in development: The participants of the women's group were a bit more critical in regard to the available functions and also were more confident in putting the robots into operation after having gotten to know them.

Robot performance

Anki Vektor was basically perceived more as a toy than as a support in everyday life and therefore dropped from the further development process later in the RoboGen project. The seniors gave several recommendations for the further development of an amended version of the Q.bo One and its combination with Echo Show features. This applies in particular to physical limitations that are common among seniors, such as problems with hearing (e.g. Ciorba et al., 2012), vision (e.g. Chen & Thomas, 2010) or functional body limitations (e.g. Mullen et al., 2012; Paterson & Warburton, 2010).

Thus, it is important for them to add a function that allows the selection of a voice according to their personal preferences so that the robot's language is understandable and sounds pleasant to them. The further functionalities that are expected to make the systems easy to use include simplified voice control (wake-up word) and adaptability of voice pitch and volume. Furthermore, it should be possible to change and enlarge the font size for Echo Show. Purely voice-controlled decision trees should work with short text passages from the robot in their selections and only 2-3 options should be offered so that the users just have to remember short texts. When it comes to physical limitations like difficulty in stooping, kneeling, or walking the seniors regard the use of a mobile robot as an added value.

The participants' recognition of emotions from facial expressions requires a higher accuracy than was given in the tests, for this to be integrated as a useful function in Q.bo One. The further development of the robot will therefore need to critically reflect on frequently reported biases affecting accuracy of classifying the genders or ethnicities (e.g. Buolamwini & Gebru, 2018; Howard & Borenstein, 2018). The participants' further wish was also to combine selected, recognized emotions with an action or a tip.

Asked for the benefits a robot might have, the seniors stated that the possibility to store contacts in the systems, which will be notified in case of an emergency, or storing and retrieving health data might bring added value for them as well as facilitate everyday life like for example calling a specific person by voice command, reminders for activities, tips on recipes, events in their area, TV program or similar. Finally, the participants were aware of the importance of data protection and their discussions go in the direction that there must be concessions if, for example, functions for emergencies are desired from the devices. They all agreed that it must always be possible to end the visibility to the outside world by simply pressing the "off" button.

Above all, what we learned from our research with involving potential end users in relation to gender-sensitive design is that gender is rarely directly addressed, but circumscribed: It is hidden in statements on the appearance design, voice of robot, name of the robot, tasks the robot is performing etc. Especially with older adults we experienced dominant binary models of gender in their thinking. We see it as our responsibility as researchers to carefully handle these mental models by providing several options of voices (e.g., also a gender-neutral voice), names, tasks the robot performs. Another option would be to be even consciously disruptive and use for instance male gender markers with typical female tasks to avoid perpetuating stereotypes.

Gender Awareness and Implementation Challenges

By maintaining constant gender monitoring and practicing reflexivity within the RoboGen project, we were able to identify and address our own biases and "blind spots" regarding gender stereotypes. This approach aligns with the common practice in the HCI research community (Rode, 2011) and has facilitated the integration of gender inclusion in all our analysis activities. However, as our project also showed, being aware of the problem space is essential, but it does not guarantee that we have the solutions to address these issues effectively. Recognizing the importance of presenting design and interaction alternatives that embrace the diversity of our target group is just one aspect of the challenge. Implementing these alternatives into technology poses what is known as a "wicked problem" (Rittel & Webber, 1973) — a problem that resists complete definition and resolution. It is important to acknowledge that the technological solutions we create, even with the best of intentions, are context-specific and appropriate only for their target groups, particularly in relation to gender sensitivity, as gender is a dynamic and fluid concept.

In other words, all of us involved in social robot development projects must continually educate ourselves about gender and stay updated on the latest theories to prevent the perpetuation of stereotypes. This ongoing learning process is crucial to ensure that our technology remains responsive to the diverse needs and experiences of individuals across the gender spectrum.

While it could be cautiously assumed that a lower level of technical development in a robot helps avoid gender stereotypes, this assumption might not even fully hold true for the systems used in our workshops and potentially not even for our final prototype robot "Chris" - the Q.Bo One robot enhanced with the Echo Show. Our research shows that the gender-sensitive design of social robots should not solely focus on getting the appearance design "right". Instead, it should be viewed as an integral part of the social construction processes surrounding gender, wherein stereotypical assumptions are directly reflected in the artifact design. Therefore, it remains essential to approach gender considerations in a theoretically reflective and methodologically systematic manner to address problematic gendering effectively (Ernst & Horwath, 2014; p.10).

The initial stages of our development process were informed by expert interviews and focus groups, yielding valuable recommendations for enhancing gender sensitivity within our project. Building upon these recommendations, we proceeded to develop interaction scenarios for "Chris." An interaction scenario essentially depicts a story describing a specific situation where one or more individuals engage with technology. As previous research shows (Eyssel) already these scenarios can involve gender stereotypes. Therefore, we will draw upon the insights of Paulitz & Prietl (2014), who have suggested methods for further refining the concept to integrate perspectives of social inequality and continuously reflect on the potential pitfalls arising from biased assumptions (also see Bath, 2009; J. Weber & Bath, 2007).

ACKNOWLEDGMENTS

We greatly acknowledge the financial support from the Austrian Federal Ministry for Climate Protection, Environment, Energy, Mobility, Innovation and Technology (BMK) within the 6th Call "FEMtech Forschungsprojekte". Project Number: 866694 – Gender Sensitive Interaction with Social Robots and from the Austrian Science Foundation (FWF) under grant agreement No. V587-G29 (SharedSpace).

REFERENCES

Ahmadi, M., Eilert, R., Gäckle, K., & Marsden, N. (2018). *Gender als Faktor bei der partizipativen Softwaregestaltung in Living Labs*. https://doi.org/10.18420/MUC2018-WS02-0448

Akrich, M. (1992). The De-Scription of Technical Objects. In W. E. Bijker & J. Law (Eds.), *Shaping Technology/Buildung Society: Studies in Sociotechnocal Change: General Introduction* (pp. 205–224). MIT Press. http://www.conceptlab.com/notes/akrich-1992-description-technical-objects.html

Alexander, E., Bank, J., Yang, J. J., Hayes, B., & Scasellati, B. (2014). *Asking for help from a gendered Robot* (P. Bello & Cognitive Science Society, Eds.). Curran.

Bardzell, S. (2010). Feminist HCI: Taking stock and outlining an agenda for design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10), 1301– 1310. DOI: https://doi.org/10.1145/1753326.1753521

Bath, C. (2009). *De-Gendering informatischer Artefakte: Grundlagen einer kritisch-feministischen Technikgestaltung* [Universität Bremen]. https://media.suub.uni-bremen.de/bitstream/elib/360/1/00102741-1.pdf

Brenssell, A., & Lutz-Kluge, A. (2020). Partizipative Forschung und Gender: Emanzipatorische Forschungsansätze weiterdenken.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, *81*, 1–15. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Butler, J. (1991). *Das Unbehagen der Geschlechter*. Suhrkamp.

Butler, J. (2004). *Undoing Gender*. Routledge.

Carpenter, J., Davis, J. M., Erwin-Stewart, N., Lee, T. R., Bransford, J. D., & Vye, N. (2009). Gender Representation and Humanoid Robots Designed for Domestic Use. *International Journal of Social Robotics*, *1*(3), 261–265. https://doi.org/10.1007/s12369-009-0016-4

Cearley, D. W. (2020). *Gartner Top 10 Strategic Technology Trends for 2020*. Gartner INC. https://emtemp.gcom.cloud/ngw/globalassets/en/publications/documents/top-tech-trends-2020-ebook.pdf

Chen, Y. A., & Thomas, M. (2010). Vision Screening in the Elderly: Current Literature and Recommendations. *University of Toronto Medical Journal*, *87*(3), 166–169. https://doi.org/10.5015/utmj.v87i3.1237

Ciorba, A., Bianchini, C., Pelucchi, S., & Pastore, A. (2012). The impact of hearing loss on the quality of life of elderly adults. *Clinical Interventions in Aging*, 159. https://doi.org/10.2147/CIA.S26059

Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory, and Antiracist Politics. *The University of Chicago Legal Forum*, 139–167.

Daugherty, P. R., & Wilson, H. J. (2018). *Human + Machine: IReimagining Work in the Age of AI*. Harvard Business Review Press 2018.

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 679–704. https://doi.org/10.1098/rstb.2006.2004

Dufour, F., & Ehrwein Nihan, C. (2016). Do robots need to be stereotyped? Technical characteristics as a moderator of gender stereotyping. *Social Sciences*, *5*(3), 27.

EIGE European Institute for Gender-Equality. (2018). *ICT specialists by sex (Source: Eurostat from 2016-06-07)*. https://eige.europa.eu/gender-statistics/dgs/indicator/ta_wrklab_lab_employ_selected_kwnd__isoc_sks_itsps/datatable

Ernst, W., & Horwath, I. (Eds.). (2014). Gender in science and technology: Interdisciplinary approaches. Transcript Verlag.

Eurostat. (2020a). *Population Structure and Ageing*. Statistcs Explained. ISSN 2443-8219

Eurostat. (2020b). *Mortality and Life Expectancy Statistics*. Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Mortality_and_life_expectancy_statistics#Life_expectancy_increased_in_EU-27_in_2018

Eyssel, F., & Hegel, F. (2012). (S)he's Got the Look: Gender Stereotyping of Robots1: GENDER STEREOTYPING OF ROBOTS. *Journal of Applied Social Psychology*, *42*(9), 2213–2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

Gildemeister, R. (2010). Doing Gender: Soziale Praktiken der Geschlechtsunterscheidung. In R. Becker & B. Kortendieck (Eds.), *Handbuch der Frauen und Geschlechterforschung* (3rd ed., Vol. 35, pp. 137–145). VS Verlag.

Gildemeister, R., & Wetterer, A. (1992). Wie Geschlechter gemacht werden. Die soziale Konstruktion der Zweigeschlechtlichkeit und ihre Reifizierung in der

Frauenforschung. In G. Axeli-Knapp & A. Wetterer (Eds.), *TraditionenBrüche Entwicklungslinien feministischer Theorie* (pp. 201–254). Kore.

Hartung, S., Wihofszky, P., & Wright, Michael. T. (Eds.). (2020). *Partizipative Forschung: Ein Forschungsansatz für Gesundheit und seine Methoden*. Springer VS.

Howard, A., & Borenstein, J. (2018). The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity. *Science and Engineering Ethics*, *24*(5), 1521–1536. https://doi.org/10.1007/s11948-017-9975-2

Kuchenbrandt, D., Häring, M., Eichberg, J., & Eyssel, F. (2012). Keep an Eye on the Task! How Gender Typicality of Tasks Influence Human–Robot Interactions. In S. S. Ge, O. Khatib, J.-J. Cabibihan, R. Simmons, & M.-A. Williams (Eds.), *Social Robotics* (Vol. 7621, pp. 448–457). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34103-8_45

Lee, J.-E. R., & Nass, C. I. (2010). Trust in computers: The computers-are-social-actors (CASA) paradigm and trustworthiness perception in human-computer communication. *Trust and Technology in a Ubiquitous Modern Environment: Theoretical and Methodological Perspectives*, 1–15. https://doi.org/10.4018/978-1-61520-901-9.ch001

Lee, H. R., Tan, H., & Šabanović, S. (2016, August). That robot is not for me: Addressing stereotypes of aging in assistive robot design. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 312-317). IEEE.

Lee, H.R., Cheon EJ, de Graaf, M., Alves-Oliveira, P., Zaga, C., and Young, J. (2019). Robots for social good: Exploring critical design for HRI. In 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI '19), 681-682. DOI: https://doi.org/10.1109/HRI.2019.8673130

Lee, H. R., Cheon, E., Lim, C., & Fischer, K. (2022, March). Configuring humans: What roles humans play in hri research. In 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 478-492). IEEE.

Lenz, I. (2010). Intersektionalität: Zum Wechselverhältnis von Geschlecht und sozialer Ungleichheit. In R. Becker & B. Kortendieck (Eds.), *Handbuch Frauen- und Geschlechterforschung: Theorie, Methoden, Empirie* (3rd ed., pp. 158–165). VS Verlag.

Mayring, P. (1990). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (2nd ed.). Deutscher Studienverlag.

Mayring, P. (2002). *Qualitative Sozialforschung* (5th ed.). Beltz Verlag.

Messmer, R., & Schmitz, S. (2004). Gender Demands on eLearning. In K. Morgan, C. A. Brebbia, J. Sanches, & A. Voiskuonsky (Eds.), *Human Perspectives in the Internet Society: Culture, Psychology, and Gender* (4th ed., pp. 245–254). WIT-Press.

Michel, J.-P., & Ecarnot, F. (2020). The shortage of skilled workers in Europe: Its impact on geriatric medicine. *European Geriatric Medicine*, *11*(3), 345–347. https://doi.org/10.1007/s41999-020-00323-0

Mullen, S. P., McAuley, E., Satariano, W. A., Kealey, M., & Prohaska, T. R. (2012). Physical Activity and Functional Limitations in Older Adults: The Influence of Self-Efficacy and Functional Performance. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *67B*(3), 354–361. https://doi.org/10.1093/geronb/gbs036

Mutlu, B., Osman, S., Forlizzi, J., Hodgins, J., & Kiesler, S. (2006). *Task Structure and User Attributes as Elements of Human-Robot Interaction Design*. 74–79. https://doi.org/10.1109/ROMAN.2006.314397

Neven, L. (2010). 'But obviously not for me': robots, laboratories and the defiant identity of elder test users. *Sociology of health & illness*, *32*(2), 335-347.

Nomura, T. (2017a). Robots and Gender. *Gender and the Genome*, *1*(1), 18–25. https://doi.org/10.1089/gg.2016.29002.nom

Nomura, T. (2017b). Chapter 47 - Robots and Gender. In Principles of Gender-Specific Medicine (Third Edition), Marianne J. Legato (ed.). Academic Press, San Diego, 695-703. DOI: https://doi.org/10.1016/B978-0-12-803506-1.00042-5

Ogbonnaya-Ogburu, I.F., Smith, A. DR. To, A., and Toyama, K. (2020). Critical race theory for HCI. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20), 1–16. DOI: https://doi.org/10.1145/3313831.3376392

Paterson, D. H., & Warburton, D. E. (2010). Physical activity and functional limitations in older adults: A systematic review related to Canada's Physical Activity Guidelines. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 38. https://doi.org/10.1186/1479-5868-7-38

Paulitz, T., & Prietl, B. (2014). Geschlechter- und intersektionalitätskritische Perspektiven auf Konzepte der Softwaregestaltung. In N. Marsden & U. Kempf (Eds.), *Gender-UseITHCI, Usability und UX unter Gendergesichtspunkten* (pp. 79–89). De Gruyter Oldenbourg. https://doi.org/10.1515/9783110363227.53

Robertson, J. (2010). Gendering humanoid robots: Robo-sexism in Japan. Body & Society. DOI: https://doi.org/10.1177/1357034X10364767

Robertson, J. (2017). Robo Sapiens Japanicus: Robots, Gender, Family, and the Japanese Nation. Univ of California Press, Oakland, CA.

Rode, J.A. (2011.) Reflexivity in digital anthropology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '11), 123-132. DOI: https://doi.org/10.1145/1978942.1978961

Rittel, H. W., & Webber, M. M. (1973). 2.3 planning problems are wicked. *Polity*, *4*(155), e169.

Roff, H,M. (2016). Gendering a Warbot. International Feminist Journal of Politics 18, 1, 1-18. DOI: https://doi.org/10.1080/14616742.2015.1094246

Schiebinger, L. (2019). The Robots are Coming! But should they be gendered? *Awis Magazine*, *feature*(winter), 18–22.

Schlesinger, A., Edwards, W.K., and Grinter, R.E. (2017) Intersectional HCI: Engaging identity through gender, race, and class. In Proceedings of the 2017 CHI

Conference on Human Factors in Computing Systems (CHI '17), 5412–5427. DOI: https://doi.org/10.1145/3025453.3025766

Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive Robotics: The influence of robot gender on human behavior. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2563–2568. https://doi.org/10.1109/IROS.2009.5354116

Strait, M., Briggs, P., & Scheutz, M. (2015). Gender, more so than Age, Modulates Positive Perceptions of Language-Based Human-Robot Interactions. In M. Salem, A. Weiss, P. Baxter, & K. Dautenhahn (Eds.), *4 th International Symposium on New Frontiers in Human-Robot Interaction* (pp. 18–25). https://hrilab.tufts.edu/publications/straitetal15aisb.pdf

Stroessner, S. J., & Benitez, J. (2019). The Social Perception of Humanoid and Non-Humanoid Robots: Effects of Gendered and Machinelike Features. *International Journal of Social Robotics*, *11*(2), 305–315. https://doi.org/10.1007/s12369-018-0502-7

Tay, B., Jung, Y., & Park, T. (2014). When stereotypes meet robots: The double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior*, *38*, 75–84. https://doi.org/10.1016/j.chb.2014.05.014

Wajcman, J. (2000). Reflections on Gender and Technology Studies. *Social Studies of Science*, *30*(3), 447–464.

Wajcman, J. (2004). *TechnoFeminism*. Polity Press.

Wajcman, J. (2010). Feminist theories of technology. *Cambridge Journal of Economics*, *34*, 143–152.

Walgenbach, K. (2012). *Intersektionalität—Eine Einführung.* www.portal-intersektionalität.de

Walgenbach, K. (2013). Postscriptum: Intersektionalität—Offenheit, interne Kontroversen und Komplexität als Ressourcen eines gemeinsamen Orientierungsrahmens. In H. Lutz, M. T. Herrera Vivar, & L. Supik (Eds.), *Fokus Intersektionalität: Bewegungen und Verortungen eines vielschichtigen Konzeptes* (2nd ed., pp. 265–279). Springer VS.

Weber, J., & Bath, C. (2007). 'social' Robots & 'Emotional' Software Agents: Gendering Processes and De-Gendering Strategies for 'Technologies in the Making.' In I. Zorn, S. Maass, E. Rommes, C. Schirmer, & H. Schelhowe (Eds.), *Gender Designs IT* (pp. 53–63). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90295-1_3

Weber, K. (2017). Demografie, Technik, Ethik: Methoden der normativen Gestaltung technisch gestützter Pflege. *Pflege & Gesellschaft*, *22*, 338–352. https://doi.org/10.3262/P&G1704338

Weiss, A., Pillinger, A., Spiel, K., & Zauchner-Studnicka, S. (2020, April). Inconsequential appearances: An analysis of anthropomorphic language in voice assistant forums. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-7).

Weiss, A., Pillinger, A., & Tsiourti, C. (2021, August). Merely a conventional 'diffusion'problem? On the adoption process of anki vector. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)* (pp. 712-719). IEEE.

Weiss, A., & Spiel, K. (2022). Robots beyond Science Fiction: mutual learning in human–robot interaction on the way to participatory approaches. AI & society, 37(2), 501-515.

West, C., & Zimmermann, D. H. (1987). Doing Gender. *Gender and Society*, *1*(2), 125–151.

Wetterer, A. (2010). Konstruktion von Geschlecht: Reproduktionsweisen der Zweigeschlechtlichkeit. In R. Becker & B. Kortendieck (Eds.), *Handbuch der Frauen und Geschlechterforschung* (3rd ed., Vol. 35, pp. 126–136). VS Verlag.