Designing Multi-Modal Conversational Agents for the Kitchen with Older Adults: A Participatory Design Study

Sanna Kuoppamäki[1] · Razan Jaberibraheem[2] · Mikaela Hellstrand[1] · Donald McMillan[2]

Abstract

Conversational agents (CA) are increasingly used to manage and coordinate household chores and everyday activities at home. However, these technologies should be adaptive to age-specific characteristics in order to be considered beneficial for the ageing population. This study presents a participatory design of a conversational agent to provide cognitive support in recipe following and nutrition advice for adults aged 65 and over. Through a qualitative thematic analysis, the study explores older adults' expectations, interactions and experiences with the agent in order to identify age-specific challenges of interacting with CAs. Data consists of a participatory design workshop with eight older adults (aged 65 and over), followed by a Wizard of Oz study with ten older adults interacting with the agent in the kitchen environment in a laboratory setting. Results demonstrate that older adults consider conversational agents as beneficial for providing personalised recipe recommendations, advising the user to choose appropriate ingredients and reminding them of their dietary intake. When interacting with the agent older adults displayed challenges with confirmation and repetition, questioning and correcting, the lack of conversational responses, and difficulties in hearing and understanding the multi-modal interaction. Older adults experience agents as collaborators, but not as conversational partners. The study concludes that the accessibility and inclusiveness of conversational agents regarding voice interaction could be improved by further developing participatory methods with older adults.

Keywords Conversational agents · Older adults · Kitchen technologies · Voice interaction · Gaze

1 Introduction

Conversational agents (CA) based on speech recognition and dialogue models are increasingly used to manage and coordinate everyday activities at home. They provide consultation, support and assistance in everyday tasks, health behaviours and information needs [1]. Research in this area has rapidly evolved, focusing mostly on the usage patterns and interactions in everyday life [2, 3]. Older adults, those aged 65 and over, have become an important user group for these technologies, because they can facilitate social interaction, support in care-related tasks [4, 5], and manage and coordinate day-to-day activities such as information seeking and entertainment [6].

The use and interaction with conversational agents can differ between age groups, and older adults may have age-specific preferences for using a conversational interfaces with respect to the purpose and effects of use [4, 5, 7]. Therefore, conversational agents need to be be adaptive to age-specific characteristics in order to be considered beneficial for the ageing population [8, 9]. Participatory design that takes participants' self-identified issues and concerns as a starting point for developing conversational agents has been proposed as a solution to design more appropriate applications for older adults [10, 11].

Against this background, this study presents a participatory design of a conversational agent for the provision of cognitive assistance in the specific context of a kitchen

✉ Sanna Kuoppamäki
sannaku@kth.se

Razan Jaberibraheem
razan@dsv.su.se

Mikaela Hellstrand
mihellst@kth.se

Donald McMillan
donald.mcmillan@dsv.su.se

[1] Department of Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, Ha¨lsov¨agen 11, 141 52 Huddinge, Sweden

[2] Department of Computer and System Sciences, Stockholm University, Borgarfjordsgatan 12, 164 55 Kista, Sweden
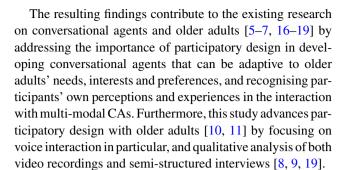
environment. Our study explores older adults' expectations, interactions and experiences of interacting with the agent in a laboratory setting where they prepared a meal in the presence of a multi-modal conversational agent. Home cooking is a central skill for independent living that ranges from nutrition and dietary choices, to knowledge and the ability to follow a recipe. Kitchen technologies designed for older adults have previously focused on smart kitchen systems for adults with cognitive impairments [12, 13], or technologies facilitating social connectedness through cooking together over distance [14]. Conversational agents that support the physical, cognitive and social aspects of this everyday task have significant potential to enhance older adults' independent living and support ageing in place [15].

Prior studies have shown that older adults perceive conversational interfaces as easier to use and learn in comparison to traditional computing devices, which indicates that conversational interfaces can improve the accessibility of digital technology [5–7, 16–18]. Older adults use conversational agents most commonly to seek information related to topics of health, local businesses, and food and drink [6]. Research has covered technical requirements for voice-based interaction with older adults [7], but less is known about older adults' own preferences for certain types of conversation, such as what kind of information older adults perceive as suitable topics and in what way an agent should approach them.

We conducted a participatory design workshop with eight older adults (aged 65 and over). The outcomes of the participatory design were taken into consideration in the development of an interaction paradigm for a kitchen focused conversational agent. This paradigm was explored through a Wizard of Oz study in a laboratory environment with ten participants from the same age group, who interacted with the agent while preparing a meal in the kitchen. We explored our participants' perceptions and responses to these design choices with video recordings of their interactions with the systems and semi-structured interviews after the task had been completed. A qualitative thematic analysis was used to identify interaction challenges characteristic for older adults: confirming and repetition, questioning and correcting, lack conversational responses, and difficulties in hearing and understanding. This empirical work was focused on answering the following questions:

RQ1 How do older adults perceive the functionality of a conversational agent providing assistance in the kitchen?
RQ2 What are the challenges of interacting with a conversational agent in a specific task that is cooking a meal?

RQ3 How do older adults experience the multi-modal interaction and design choices with a conversational agent in the kitchen?

The resulting findings contribute to the existing research on conversational agents and older adults [5–7, 16–19] by addressing the importance of participatory design in developing conversational agents that can be adaptive to older adults' needs, interests and preferences, and recognising participants' own perceptions and experiences in the interaction with multi-modal CAs. Furthermore, this study advances participatory design with older adults [10, 11] by focusing on voice interaction in particular, and qualitative analysis of both video recordings and semi-structured interviews [8, 9, 19].

The paper begins by presenting previous research on conversational agents for older adults and in home cooking. We then present the methodologies employed in this work, namely participatory design, Wizard of Oz studies, and semi-structured interviews. The results are presented in three sections, focusing on older adults expectations, interaction challenges and experiences when interacting with a conversational agent. In the discussion, we examine the findings in relation to social connectedness, multimodality, and age-specificity of older adults' interaction with a conversational agent.

2 Background

2.1 Conversational Agents and Older Adults

Research on the use of conversational agents (CA) based on speech recognition and dialogue models has rapidly evolved, focusing mostly on the usage patterns, features and interactions in everyday life [2, 3] and health care environment [1, 20–22]. Conversational agents have the potential to enhance physical, cognitive and social interactions among older adults, and support the maintenance of an active and healthy lifestyle [5, 18, 23]. Research in the field of Human—Computer interaction has investigated conversational agents among adults aged 65 and over. These studies have focused on the acceptability of anthropomorphic designs [6, 18], usability evaluation [21], accessibility of voice interfaces [24], usage patterns and satisfaction of use [25], new embodiments of conversational agents in well-being [26, 27], design requirements [4, 23] and the application of conversational agents in caregiving [5].

Older adults may experience unique challenges with such interfaces, but also unique benefits from the use of conversational user interfaces. This indicates that speech-based interaction should consider questions of synthesis choices and conversation content when supporting this age group [18]. On the one hand, using voice as an interaction paradigm can help to overcome many difficulties older adults typically experience with interactive technologies. For instance, older adults with vision and mobility impairments can leverage the conversational interfaces to complete tasks with increasing

complexity [24]. On the other hand, older adults may perceive it difficult to know how (and what) to speak to a conversational agent [18]. Accessibility issues are especially prevalent for older adults with hearing impairments [24]. Therefore, the linguistic content and default voices that conversational agents provide when interacting with older adults should be designed appropriately from the conversational user experience, and to support the more inclusive interaction.

The perceptions and motivations to use conversational agents in everyday life can also be different for older adults. Whereas children perceive agents like a person [28], adults see it as a utility that can evoke feelings of independence and empowerment [29]. Chung et al. [25] explored the unique features of older adults' interaction with a conversational agent in relation to the most frequently used functions and what they found to be satisfying and unsatisfying aspects of the agent. They reported that older adults tended to personify the agent more, indicated by the use of polite words such as 'grateful', while younger adults tended to consider it a tool and placed more importance on its convenience; older adults also perceived the music function as having a higher importance compared to the younger adults. Studies have also shown that older adults use voice assistants for specific purposes. In a study by Pradhan et al. [17], older adults used the conversational agent in information seeking tasks around topics of health, local businesses, and food and drink. Most older adults perceived the voice-interface as easier to use and learn when compared to traditional computing devices, indicating that voice-based interfaces can improve the accessibility of digital technology. They report that older adults were concerned about the reliability of reminders set with a conversational agent, in that they may forget to set the correct reminder without visual representations of the current reminders, and that unstable internet connectivity may result in reminders not being triggered. Other challenges in the use of agents included unpredictability and instability (e.g., devices timed out before completion of voice commands) and inconsistency and lack of clarity in relation to the formulation of voice commands.

Several attempts have aimed to leverage conversational agents to enhance well-being and social connectivity among older adults. Simpson et al. [26] developed 'Daisy', a voice-controlled conversational agent embodied as a household potted flower, providing companionship in times of social isolation by engaging with them in casual conversation, suggesting relevant activities to keep them connected with their community and having them care for it. El Kamali et al. [27] designed 'Nestore', a virtual coach developed to support older adults' well-being, based on two different types of conversational interaction: a text-based chatbot integrated in a mobile application and a coach, embodied as a physical object based on vocal interaction. Zubatu et al. [5] focused

on the possibilities of conversational agents in empowering older adults with mild cognitive impairment and their care partners. Agents can support coordination and planning between older adults and their caregivers, and can amplify the support that the caregiver needs to provide. However, the utility of the conversational agents in everyday live is largely depended on how much the caregiver scaffolded the available functionality, meaning that they were responsible for setting it up and contextualising the abilities of the agent for the specific needs and desires of the users.

Usage patterns, features and interactions with conversational agents can show age-specificity, both in relation to age-specific preferences of using a conversational interface [7] and the purpose and effects of use [4, 5]. Gollash and Weber [7] identified age-specific strategies in dialogue systems and speech recognition accuracy. To respond to the needs of older adults, conversational agents should be able to correctly recognise 'unusual' formulations; complex dialogues comprising multiple pieces of information should be presented as simple or guided dialogues; agent should ask only one question per dialogue with a limited set of possible answers; it should be able to keep information about the conversation context; and so on. Furthermore, Nikitina et al. [4] describe the age-specific requirements and early system design for a smart conversational agent that can assist older adults in the reminiscence process. The practice of reminiscence has well documented benefits for the mental, social and emotional well-being of older adults. However, the technology support is still limited in terms of need of human presence, data collection capabilities, and ability to support sustained engagement, thus missing key opportunities to improve care practices, facilitate social interactions, and bring the reminiscence practice closer to those with less opportunities to engage in sessions with a trained companion.

2.2 Conversational Agents in Home Cooking

Technological support designed for kitchen activities has traditionally focused on instructional guidance and multi-modal feedback on cooking, such as following recipes in the correct order. Hamada et al. [30] used multimedia to offer instructional guidance on cooking based on the interpretation of cooking workflows and rescheduling recipe steps. Doman et al. [31] offered instruction for each cooking step by means of personalised videos. Sato et al. [32] designed a "MimiCook" that displayed written instructions with video-projection to offer immediate feedback for the person in the kitchen. Recent developments in the kitchen also include voice assistants and machine learning applications with the ability to understand and predict user's information needs [33]. For instance, Lim et al. [34] used multi-modal machine learning in order to develop effective

recipe recommendations and minimise the overhead of tracking kitchen ingredients.

In instructional 'how-to' videos that are commonly used in learning new recipes in cooking, users often need to clarify misunderstandings, skip familiar contents, or jump to the later parts to see the result and prepare for future steps. Adopting a conversational user interface can decrease this cognitive load when people can control the video with voice while performing the tasks with hands [35]. Chang et al. [36] presented 'RubySlippers', a system that supports efficient content-based voice navigation through keyword-based queries, through which users could perform tasks with fewer commands and less frustration than in the conventional voice-enabled video interface. Different navigation objectives include pace control pause, content alignment pause, video control pause, reference jump, replay jump, skip jump, and peek jump. Designing voice-based navigation for how-to videos should therefore support conversational strategies like sequence expansions and command queues, that allow users to identify and refine their navigation objectives explicitly, and support the seven interaction intents [35].

To date, little or no research has investigated how conversational agents can facilitate cognitive support among older adults in domestic tasks that contain complexity in organisation, managing and synchronising tasks. While cooking has been recognised as an important communal activity and skill for independent living for persons with cognitive impairments residing in sheltered living facilities [12, 13], these studies have been limited to auditory instructions displaying cooking steps with verbal exchange such as using the voice of the caretaker, or video instructions where the caretakers showed how to accomplish the next cooking step [13]. Our study aims to investigate conversational agents for cognitive support based on in-depth analysis of older adults' interactions and experiences to ensure that such technologies reinforce their own skills and competences and encourage their participation [15].

3 Data and Methods

3.1 Participatory Design Approach

Participatory design (PD) has been offered as a solution to overcome challenges in the design process of conversational agents for older adults [10]. These approaches take participants' self-identified issues and concerns as a starting point for developing potential applications for conversational agents, based participants' interpretation of the capabilities of these systems. Participants are given a possibility for critical discussions of the potential social consequences and meanings of CAs [11]. Participatory design is expected to result in mutual learning between participants, and promote active

participation of older adults as designers, rather than only users of conversational systems [10]. However, participatory design with older adults often entails challenges, and should always be based on building trust and mutual understanding of the everyday life condition among older adults. Participants may not have experience with the conversational technologies, and they may not see themselves as designers of any technology [37]. Therefore, our PD approach employed both focus group discussion and design scenarios as well as qualitative evaluation of the usability and acceptability of the agent [19, 38–42].

3.2 Participatory Design Workshop

We conducted a participatory design workshop with older adults (aged 65 and over) to investigate their expectations towards functionalities of conversational agents in the kitchen environment (RQ1). Older adults' self-identified concerns and interpretations were taken as a starting point for the design [10] in order to learn how they interpret existing conversational agents, and also introduce two conversational platforms to our participants. We aimed to understand issues influencing the acceptability of conversational agents for our participants in order to develop an interaction paradigm that could be perceived as purposeful among this age group [5, 6, 16–18].

The workshop consisted of three stages: (a) presentation of potential conversational agent interactions with images, (b) discussion, ideation and visualisation through the co-development of design scenarios, and (c) a group discussion prompted by the demonstration of prototype conversational agents in cooking-related activities.

In the first stage, we began by presenting two conversational platforms providing a speech-based instruction for three different activities: tool suggestion, recipe reminding and orientation of action in the kitchen. These were based on results from our previous video study on older adults' cooking at home [15]. These scenarios also presented conversational agents that employed two different linguistic styles: direct commands and conversational suggestions. This prompted a focus group discussions in which we asked participants about their initial impressions of the agent, the linguistic style and content, and the acceptability such an agent in kitchen-related tasks in their own everyday lives.

After the priming focus group discussion, the participatory design session started with a demonstration of live conversational assistance, where researchers demonstrated the use of conversational agent in the kitchen using a virtual agent on a screen. Participants then worked together to design and document conversational scripts they would perceive purposeful with the agent, discussing benefits and drawbacks uncovered in the earlier discussion and virtual agent demonstration.

The final phase involved the researchers role-playing a selection of the interactions designed by the participants, using the screen-based virtual agent and a wizarding interface. During these demonstrations, the participants were encouraged to discuss the interaction as it progressed, suggest changes and solutions to problems that arose, and share their initial observations and concerns.

The participatory design workshop was audio recorded, and all discussions between participants and researchers were transcribed verbatim.

3.3 Wizard of Oz Study

To explore older adults' interaction challenges with a conversational agent (RQ2), we developed an interaction paradigm that was tested using a Wizard of Oz trial facilitated by the Tama platform. Tama is a conversational agent that can use gaze as well as voice to provide multi-modal interaction [43]. The interaction was tested through having individual participants follow a recipe with the support of the physical Tama smart-speaker on the counter-top. The system produced voice utterances when directed by a researcher via custom web interface, using the Google text-to-speech service in Swedish (the native language of our participants). For half of the participants, the Tama platform was used with the mutual-gaze system activated. For the other half of the participants, it was deactivated. This meant that it would look at the participants while it was speaking to them, and if they looked at the device it would look back and indicate to the Wizard that the user had initiated interaction. When using automated responses, the mutual gaze system can be used instead of the standard 'wake-word' (e.g. 'Ok, Google' or 'Hey, Siri') in conversational agent interaction—removing the need to start each interaction with the name of the device. With respect to initiating interactions, in this wizarded scenario the agent responded with or without the participant using 'Tama' as a wake-word or achieve mutual gaze with the system. In this way, in the conditions where the gaze was active it was primarily used to visually indicate that the users' utterances had been received or that (in the proactive interactions described below) the device has information to impart to the user.

The interaction paradigm instructed the wizard in what they could and could not reply to, and how the replies should be formulated if they fell outside the pre-programmed recipe and timer-based utterances. We describe this paradigm below through the task vocabulary, recipe progression, and proactive alternatives available.

3.3.1 Task Vocabulary

First, we developed a shared vocabulary within the bounds of the cooking interaction drawing directly from the textual description of the recipe. The wizard interface was populated with both the steps of the recipe and the list of ingredients with their quantities (where specified in the recipe) to be triggered with a single click. To simulate a shared context between the user and the agent regarding the recipe, the wizard replied to 'How many' or 'How much' questions by instructing the agent to read out the related ingredient's list item, and to other queries regarding the ingredients by triggering the text-to-speech response of the closest recipe instruction.

Beyond the list of ingredients, the list of instructions was also considered as a resource for shared linguistic context allowing user utterances to be keyword- matched to instructions, that would then be triggered to be read to the user. Instruction list items which included specific lengths of time generated a pre- configured timer next to that instructions. These were also seen as objects in the context of the ongoing task to be queried by the user.

3.3.2 Recipe Progression

In recipe progression, the agent was conceptualised as being able to comprehend the ongoing plan within the sequence of the recipe, and the user's progress through those steps. This included the ability to track ingredients listed in the recipe when they were either on the chopping board or added to the pot. This enabled the illusion that CA could respond to direct questions about the current state of the ongoing cooking action.

3.3.3 Proactive Suggestions

Proactive alternatives and reactive instructions were also provided by the agent, as suggested by participants during the design workshop. These fit into the Organisational Assistance category of technological intervention [15]. This type of assistance is provided through proactively offering advice such as optional alternatives or additions to the current recipe step to adjust taste or texture to the users' preferences, as well as advice related to the impact that substitutions in the ingredient lists would have on the task. For these optional dialogues, the system would utter an interrogative 'hmmm'—and in the gaze condition, look towards the user as well—until the user indicated that they would like to receive the suggestion.

The conversational agent also proactively suggested the use of timer functionality within the ongoing interaction, where deemed appropriate by the wizard. This tended to be realised as a follow-up question to an instruction that involved a cooking stage that lasted a specific duration (e.g., an instruction including 'simmer for 3 min' would result in a follow up question of 'should I start a timer?').

3.4 Procedure

The procedure begun by negotiating informed consent with the participants for (1) participation in the study, (2) documenting the study with video recordings and (3) publishing anonymised extractions from the data.

The participants were then introduced to the interaction paradigm, the platform and the recipe they were asked to cook. The participants were informed that they were expected to follow the recipe, but could adjust the ingredients based on their dietary requirements. They were also told that the study was not an evaluation of their skills, and therefore they were encouraged to interact with the agent as often they felt comfortable.

The recipe consisted of 12 steps involving 14 ingredients. After participants had read the recipe, they were asked to cook the meal and ask guidance or reminders from the conversational agent as often as they wanted without access to the printed instructions. The study was organised in a university environment, where all ingredients and utensils needed for cooking the recipe were provided.

The Wizard of Oz study with the conversational agent was documented with dual-angle video recordings. The recordings started when participants started to cook a meal, and ended when the meal was ready. The video captured the voice commands when interacting with the agent, as well as their gestures and bodily movements within the kitchen space.

3.5 Semi-Structured Interview

To investigate older adults' experiences of interacting with a conversational agent (RQ3), we conducted a semi-structured interview for participants focusing on the acceptability and usability of conversational agents in the kitchen. The semi-structured interview template was developed based on a robot-acceptance model [44] and existing usability scales for conversational agents [45, 46]. The semi-structured interview consisted questions of ease of interacting with the agent, perceptions of user's privacy and security, trust towards the agent, and perceived social presence of the conversational agent [45]. The interview template was kept similar in all interviews, but according to the principles of qualitative interviews, we followed the template in a flexible order. Participants were encouraged to share personal insights that were important to them, and follow-up questions where asked to better understand 'how' and 'why' older adults perceived the interaction with conversational agent in a certain way. All participants attended the interview after having completing the cooking task with conversational agent. The interview lasted approximately 30 min, and all interview recordings were transcribed.

3.6 Participants

We recruited participants from the age group of 65 and over to attend all phases of the study. First, we invited the participants from our previous study [15] to attend the design workshop at our university. Eight (8) older adults in this age group participated in this first stage. Second, we recruited participants for the Wizard of Oz study by creating an email invitation that we distributed in several locations, such as using the previous panel of workshop attendees, sending the invitation to senior associations in Sweden and using formal and informal networks as a 'snowball method'. In total, ten (10) older adults in this age group volunteered to be part of the Wizard of Oz study and follow-up interview.

Over half (6) of the participants were female, and majority (8) were highly educated. Half of the older adults (5) had previous experience on interacting with conversational agents, but only (1) had owned a smart speaker such as a Google Home or Amazon Echo. The majority of the older adults (6) evaluated their cooking skills as relatively good. Almost all participants had used technologies such as mobile phones or tablets in cooking, but none of them had previous experience of conversational agents in cooking.

3.7 Analytical Approach

The study employed a qualitative content analysis approach to all materials collected in the study [47]. The analysis of video recordings was conducted in three phases. First, all video recordings were watched through to form a comprehensive understanding of the data. Second, interaction moments that showed typical interactions and common challenges among older adults were selected for further analysis. These clips were analysed in detail, using an interaction analysis approach to focus on the in the moment actions and reactions of the users with respect to the technology. From these we selected demonstrative clips of interaction with the agent to be transcribed for inclusion in the manuscript.

Interview material from the design workshop and Wizard of Oz study was transcribed. These transcriptions were analysed with thematic analysis, by coding and re-coding meaning units from the transcribed text. In the first phase, all material was read through to achieve a holistic understanding of the data. In the second phase, meaning units that respond to older adults' experiences of the multi-modal interaction were coded from the data. In the third phase, these meaning units were further categorised to separate themes. In reporting the results, participants have been given pseudonyms and their faces are blurred.

4 Results

4.1 Older Adults' Expectations Towards Conversational Agents in Home Cooking

Design scenarios presented during the workshop resulted in discussions on expectations towards conversational agents in home cooking. Older adults perceived four main functionalities for these agents: recipe following, nutrition advice, motivation for cooking and practical support.

4.1.1 Recipe Following

Older adults discussed how instructional videos for cooking could be supported by speech and voice interaction, rather than touch or screen, that can become unhandy in the kitchen. One participant, Lisa, explains the difficulties of using touch or screen as an interaction modality in the kitchen, and proposes a conversational agent with whom to discuss the recipe with. She imagines the agent to be a collaborator who advises on which tools or ingredients to use:

> I sometimes find it very difficult when you have to take a recipe for.. iPad, and it just goes away all the time..then it would be good if the robot had it inside, as well as to discuss the recipe with . (-) We follow the recipe together, and then, (the robot) says 'now you take this and this', instead of me running back to the computer and push the recipe again and holding on. Then the robot could have it inside programmed, and I do not have to keep on looking and looking at my iPad then. (Lisa)

4.1.2 Nutrition Advice

Older adults perceived nutrition decisions and dietary intake as central aspects of healthy ageing, in the form of tracking nutrition intake and receiving personalised recipe recommendations. Although nutrition advice was not presented as a specific scenario for conversational assistance, older adults themselves raised nutrition intake as a age-specific need for receiving personalised recommendations to:

> Older adults do not eat enough nutrients and you know the diet becomes too one- sided..and a robot could help a lot with. (Daniel)

Advising the user to choose appropriate ingredients, reminding of dietary intake, and optimising the usage of existing ingredients in the kitchen were considered as a conversational task based on predicting the user's needs with other information available. Asking 'what to prepare for a dinner' was a conversation that older adults imagined to have with the agent. Nutrition advice for healthy ageing could be embedded as suggestions or reminders to add or choose ingredients for the purpose of filling certain dietary intake:

> If she opens the refrigerator door (–) we should also ask the robot, 'what is it we should have for lunch today' and then (the robot) thinks about the variation so that it will be good. On the other side of the doors you have, what you have eaten before. So that you need more broccoli, more carrots now for the next meal and so on. The fat mixture with dietitians and the whole choir so that..everyone eats. (Michael)

Older adults also expressed their interest for collecting data on dietary intake, and voice-based interaction system suggesting alternative solutions for the purpose of improving nutrition choices (e.g., adding more vegetables or increasing protein intake).

4.1.3 Motivation for Cooking

Older adults identified a number of social and situational factors influencing cooking habits and motivations, which should be taken into consideration when improving context-awareness of conversational agents. Cooking needs and interests often differ between family members, household type, and social situation:

> My wife and I cook very differently, she looks in the fridge what we have and then she throws together something that is delicious..I can cook a gourmet dinner with the provision that I can start from the beginning and then go shopping. (Michael)

These individual and socially shared preferences may influence the motivations and habits of home cooking, which can be associated with the perceptions of the context-awareness of the agent. Individual differences among household members were described as including the choice of ingredients, the use of kitchen tools and appliances, and the level of planning and organisation involved in undertaking kitchen-related activities. The purpose of cooking is also situated, and different motivations were reported when in oneself or cooking for others. Cooking is therefore a socially shared practice between family members, which is influenced by routines, norms and habits in the household:

> Exactly the same dish, two people do not cook in the same way at all.. they do not use the same tool yes I mean it is very different for me.. (–) and my wife she picks out a lot and I am more..organised..I think. The purpose of cooking also depends on situational factors, such as differences in the purpose of cooking for oneself or cooking for others. (Paul)

4.1.4 Practical Support

Participants perceived conversational assistance as an external 'help' and 'support', where the purpose of the interaction is pursuing a practical goal in the kitchen. Voice interaction was considered as a transactional; 'a tool rather than a dynamic social entity' [48]:

> Maybe it could fill a small conversation (need) that way.. but, you are not ready for..especially in accommodation, maybe you do not want anyone to get involved in so much. But then I think it's more about 'can you please cut the onion or chop the onion'. (–) I see it as an auxiliary function and not a control function.(–) When you have a recipe, you still want to concentrate on the recipe (Emily).

Older adults considered the home environment as a private place where they did not want to involve a social interaction with the agent outside the immediate cooking task. Participants did not like human-like features attached to conversational agents, such as visual gestures or eye contact generated by the machine, and they were more positive towards voice interaction that displays no embodied output:

> When you think of robots, it is not.. that it should be human-like, it..seems a little strange actually. (–). Why if you are going to reconstruct a human..being, then it is better, I think, that you stick to the technical..stuff (Emily).
> Yes, you could only go by the voice; If you look at Siri or them, it's speakers that you hear them on then huh and..talk (Paul).

As a solution to overcome this tension, older adults suggested personalisation of conversational agents through a selection of their own pictures such as photos of their personal choice attached to the agent's physical output, rather than agent's capabilities for providing social or emotional responses:

> That you supplement with a picture .. we know that the more things we start to use, our abilities, the better it works, so if you both hear and see the picture. It can be..the recipe, or it can be..something too, descriptive, so that you do. . . is probably good, that you reinforce what is said..it can be with pictures or photos (Emily).

Participants mentioned language-specific characteristics, such as certain intonation typical for Swedish language, which made it difficult to understand voice interaction of the agent—even though the agent was speaking in reasonable Swedish using the Swedish Google text-to-speech services the intonation were noticeably different to that of a native speaker. Voice as an interaction modality was rather new and unfamiliar to participants, and they found it challenging to articulate their preferences for certain types of voice interaction in an accurate way (e.g., in relation to speed, tone, intonation, wording, style or content).

4.2 Older Adults' Interaction with the Conversational Agent

In this section, we focus on presenting interaction challenges and analysing the type of voice commands that older adults used when interacting with the conversational agent. We identified the following interaction challenges typical for older adults: confirming and repetition, questioning and correcting, lack of conversational responses, and difficulties in hearing and understanding.

4.2.1 Confirming and Repetition

Confirming commands and repetition of questions was an interaction challenge that was characteristic for the older adults in our study. They used repetition in both asking and responding commands with the agent. They also demonstrated interactions with the agent without clear context markers which would be challenging to support in traditional dialogue-based systems, shown as questions such as 'What should I do now, Tama', or 'What should I do after, Tama', as in the case of Alice. Figure 1 shows an example of such interaction, where Alice uses confirmation and repetition to receive relevant advice from the agent.

Figure 2 is another example of interaction pattern including repetition and asking the same question many times. Emma starts the interaction with a question regarding the amount of ingredients, which the Wizard was unable to hear and therefor the agent only responds 'Yes?' Emma asks the same question again, following a standard conversational repair tactic of speaking at a higher volume and more slowly. This time the agent repeats the corresponding recipe step—which did not include the answer to her questions. After that, Emma further modifies her question, reducing the complexity and asking 'How many?', at which point the Wizard manually added the implied number of onions which was absent from the recipe. This example shows that the need for repetitive interaction was sometimes a consequence of the agent's limited capabilities to understand or interpret utterances from the user, and that understanding how such repetition is used to repair both interactions and shared understandings of context is a valuable resource to rich conversational interaction design.

4.2.2 Questioning and Correcting

Older adults frequently adjusted, questioned or corrected the advice received from the agent. This was shown as a situational or contextual adjustment of the advice to be applicable

Fig. 1 Confirming and asking


```
Alice: What should I do now, Tama?
CA: Chop the onion and put it into the
     pan. Season with salt and pepper.
Alice: Was it one onion? (.) Tama?
       (0.6) Tama?
CA: Chop the onion into small pieces.
Alice: What should I do next?
```

Fig. 2 Repetition and asking


```
Emma: Tama how many onions should I
      take, Tama?
CA:   Yes?
Emma: How many onions should I take?
CA:   Chop the onions and put it to the
      pan. Season with salt and pepper.
Emma: How many?
CA:   One.
Emma: One onion. OK
```

Fig. 3 Rejecting suggestion


```
CA:    If you want to, I can start the
       timer.
Clara: No, no. There is no water. I have
       no water now. There should be
       water.
CA:    You can do what you want to.
Clara: I need to have water. Otherwise,
       I cannot cook.
CA:    You can do that.
Clara: Ok, I will take the water.
```

in their personal context. The adjustment was displayed in those situations where the advice received from the agent was regarded as too general, and thus irrelevant for the user. For instance, when the agent asked to take "a small garlic", participant replied "Ok, then I will take half garlic" (Clara). In Fig. 3, Clara rejects the agent's suggestion to set a timer, by mentioning that she does not have the water yet, that is be needed for the next step in the cooking process:

Figure 4 is an example of older adults questioning the advice received from the agent. James starts the interaction with the question 'Should I put the bell pepper now, Tama?'. When the agent only responses with 'Yes', James asks again 'Are you sure about it?' following up with 'Haloo, haloo?' when the agent was slow to respond. The agent replies 'Yes' again, and then James replies'OK'. This type of questioning could indicate either a lack of trust towards the agent, and in this case, older adults' familiarity with the recipe and level of cooking skills.

4.2.3 Lack of Conversational Responses

Older adults engaged in only a few conversational responses with the conversational agent outside the immediate cooking task. Older adults used rather complex queries, but they were all focused on the task at hand, rather than having a social interaction with the agent. This finding in consistent with previous studies, which show that older adults perceive conversational agents as a utility that can provoke sense of independence [29]. Lack of conversational responses was pronounced by focusing strictly on task-based utterances, and using relatively short voice interaction. For instance, after asking the agent advice on the next steps, Alice (Fig. 5) responded only with a short statement, and did not respond to the agent's proactive humming that was used as a means for stimulating interaction.

Older adults often used a polite vocabulary with the agent, such as responding with 'good' and 'thank you' for the agent's proactive suggestions. However, these polite interjections were still focused on task-completion and recipe

Fig. 4 Questioning suggestion


```
James: Should I put the bell pepper now,
       Tama?
CA:    Yes.
James: Are you sure about it?
       (0.8) Haloo?
       (0.6) Hal[ooo]?
CA:           [Yes.]
James: OK
```

Fig. 5 Lack of conversational responses


```
Alice: Tama, should I cut the onion?
CA:    Cut the onion and put it into the
       pan. Add salt and pepper.
Alice: Fine.
       (2.8)
CA:    Hmmmm
       (2.5)
       Hmmmm
```

Fig. 6 Polite vocabulary


```
Sam:   Could you start the timer?
CA:    Absolutely.
Sam:   Great, thank you.
CA:    You're welcome.
```

related interactions, rather than signalling a desire for social interaction with the agent. Figure 6 is an example of such interaction where polite interjections between the user and the agent resulted more turns of talk, but did not result in a social interaction. The ability to engage with this sort of dialogue, however, could have contributed to Sam showing more engagement with the agent, and responding more consistently to the proactive interactions than many other participants.

4.2.4 Difficulties in Hearing or Understanding

Older adults encountered frequent miscommunications with the agent, in the form of not hearing the voiced instructions, misinterpreting the utterances or asking too many questions of the agent in too short a time, faster than the Wizard and the text-to-speech system could respond to. For Emma (Fig. 7), turn-taking turned out to be challenging; she asked the same question two or more times before the agent had been able to give an answer. This could be resulting from a lack of experience of interacting with such agents, where the cadence of question and answer is slower than in human–human conversation. However, it points to an opportunity for interaction design where the agent could 'hold the floor' by indicating

that it will take its turn to talk, even before the system has been able to process the user's utterance and formulate a reply.

Figure 8 shows a typical miscommunication regarding the use of the proactive timer. Emma showed several interactions with the agent that were characterised by miscommunication due to difficulties in hearing or understanding the machine generated voice. As part of the proactivity designed into this interaction paradigm, the agent gave advance warning before a timer would finish. This was to allow the user to prepare for the next step, if that was needed. One example was users being able to unhurriedly make space to take the pot off the hob when the time was up. To do this, the agent informs the user that 'soon there will be alarm', but Emma does not understand this, and responds 'What did you say?'. The agent repeats the advance alarm warning, but Emma continues 'I do not understand'. This could also indicate that in the cases of lack of situational information from the agent, users may struggle to contextualise their advice, which may be shown as lack of understanding the agent.

Fig. 7 Difficulties in hearing

```
CA:     …3 minutes. Should I start a
        timer?
Emma:   I cannot find any timer.
CA:     Should I start a timer?
Emma:   I cannot find any timer.
CA:     I can start the timer, if you
        want me to.
Emma:   Yes, two minutes.
```

Fig. 8 Difficulties in understanding

```
CA:     Soon there will be an alarm.
        (1.2)
Emma:   What did you say?
CA:     Soon there will be an alarm.
Emma:   I don't understand.
CA:     Soon there will be an alarm.
        (6.4)
        <beeping alarm sound>
```

4.3 Older Adults' Experiences of Interacting with Conversational Agents

In this section, we explore older adults' own experiences when interacting with a conversational agent during the cooking session based on their responses during the interview, and how they consider the purposefulness and functionality of conversational agents in relation to their previous experiences of home cooking and interactive technologies.

4.3.1 Collaboration with the Agent

Most older adults considered the agent be entertaining, but they interpreted the agent to be a practical support in daily tasks, rather than a conversational partner. Older adults mentioned maintaining a clear boundary on what kind of discussion they would like to have with the agent, but considered it important for the agent to provide encouragement and social support. This resembles ideals of agents as collaborators that will help them to achieve practical tasks:

> Yes, but (–) it encourages me because now I've been good, now I've made a soup, 'good'! a bit like that..it's about the same as she does with gymnastics huh..'oh how good you've been today' and 'how nice this has gone' and so on, so..I think Tama would feel (like that).. (Maria)

This encouragement in the form of polite but relatively short social responses from the agent seemed to increase the likeability of the agent, so that the agent was considered to be kind and friendly, and able to understand their needs. Hence, older adults described the agent as a friend that can help in practical matters:

> I saw it as a, uh..friend, who is in the kitchen, and who keeps track of the recipe and helps me to figure things out. (Samuel)

4.3.2 Social Responses with the Agent

Older adults regarded the interaction with the agent to be effortless and easy, but wanted to draw a clear boundary on what kind of conversations they would be willing to have with the agent. They considered the agent to be a 'helper, but not a conversational partner'. The proactivity of the agent in the form of the use of the interrogative 'hmmm' provoked reactions of wanting to start a conversation with the agent, such as searching for additional confirmation from the agent for the right recipe steps.

> I thought it was a bit fun, like.. I found myself wanting to hold (conversation)..like this when (the robot) sighed. Do you think I'm doing something wrong or..?', or 'Is it taking too long?' (–) but, but then I think it becomes like a..conversation partner and I think that you have to..you have to be careful about. I think you have to have to draw some line (–) A robot is a robot! It's not an uh..a..conversation partner..then you get to talk to someone you trust. (Alice)

In the example above, Alice describes how the proactivity of the agent caused her to question her own actions. She interpreted the 'hmmm' from the agent as an indication that she was 'doing something wrong' or not performing the recipe steps effectively enough. This proactivity also provoked a sense of willingness to start a conversation with the agent even without a clear purpose to the interaction, which one participant found problematic. She mentioned that conversation between humans is based on mutual trust, which will not be possible with an agent. As she further explains, the social presence of the agent in the form of proactive, or polite greetings, can foster anthropomorphism of the agent and result in the user attempting to treat it as human-like in conversation, which should be avoided:

> It mustn't get too personal, (–) I had a great desire to..start a conversation..and I probably did a little mumbled probably a little there for it, like..and then it was completely quiet...but yeah, I don't know, I don't think it should be..it's clear (–) no, I think you should keep ..robot is robot anyway. (Alice)

Female participants were a bit more hesitant to see agents as conversational partners than the men. Females considered conversations with the agent on practical topics to be easier to interact with and understand when they contained some of the social characteristics of talk. Sam, for instance, describes how everyday discussions on grocery shopping with the agent could create an an experience of social interaction, even though the discussion itself is focused on practical issues:

> To help keep track of 'what's going on', 'how is the weather', (–) 'can you help me remember that I have to buy milk (—) then I have it connected via the mobile phone, so I ask Tama, 'what was it I was going to buy something for, 'well you're going to buy milk and such', and then it helps so what. (—) If it is now a helper, but at the same time has a social behavior so to speak..then you get these two things. (Samuel)

Most older adults considered that they need to adapt their own conversational style in order to fit with the agent's linguistic style. A tendency to adapt own interactions in order to avoid misunderstandings with the agent was frequently experienced among older adults. In the example below, Samuel did not expect to have a similar conversation with the agent as with humans, and inherently, he perceived a fundamental difference in human–human versus human–machine interaction:

> As a human being, I adapt a little bit in order to fit (my communication) in with it. If they (agents) should look like us humans or if they should have a different appearance. (–) To respect that it's something that

you're interacting with (–) but you shouldn't act like it's a human BUT can help you. (Samuel)

4.3.3 Miscommunication with the Agent

Video recordings showed several instances were older adults could not hear, understand or remember the voice command from the agent. These types of failures with conversational agents are common in all user groups and in various contexts [49]. Older adults considered the major reason for interaction failures to be caused by a lack of previous experience on voice interaction. They also highlighted the need of agent to understand them; indicating that older adults perceived a need to adjust their own interactions to be appropriate for the agent, such as asking questions in a different way:

> There was probably a single time she didn't understand what I said...but then I just repeated in a different way, I think. (Maria)

Hence, older adults tended to interpret problems occurring in the interaction with the agent as a result of their own interaction patterns, rather than the technical capabilities of the agent. This tendency to adjust their own actions to make sure that the agent could understand them could indicate that the fluency of interaction was associated with human, rather than tech- nical capabilities. Thus, repetitive interaction with the agent was experienced as unnatural:

> Yes but once she repeated and maybe if you're not that used to it you might feel that 'oh I'll do it again..', because it can happen sometime (–) a bit forgetful and so on, there it can be a problem..that there can be repetitions. (Maria)

Sometimes older adults regarded the proactive suggestions from the agent to be a sign that they should change their actions even though older adults themselves did not see any reason to do so. This also communicates another type of misunderstanding, where the intervention from the agent—based on certain interaction rules—could be considered as irrelevant for the user, and in some cases, provoke actions that may not be meaningful for the task completion. In these cases, older adults did not want to follow the orders from the agent, but rather decided to follow their own ways-of-doings:

> Then I already had, even then I had started chopping..in my way, the way I think you should chop an onion..but there she had a different opinion..and I thought I was right, so yes..there I continued as I wanted..like, it's also the case that it has to you know..how you want to do. (Maria)

Therefore, older adults expressed a strong need to be in control of the agent, rather than vice versa. They were comfortable in rejecting some of their suggestions and follow their own routines, whether it was perceived as 'right' or 'wrong' by the agent or not. Older adults also recognised age-specific limitation in articulating and hearing voice commands, and they were comfortable with 'not understanding' everything:

> When you get older, you also have difficulty articulating certain questions (–) because when I get really old I might slur..and then it won't understand. I used certain words on purpose, to see if it would understand that, but it.. didn't. (Oscar)

In the example below, Oscar describes a common situation of miscommunication characterised by a misalignment between older adults' expectations and the agent's commands. The participant expected the interaction go faster, and wanted to ask for a clarification. Oscar mentioned a clear reason for misunderstanding, that was not hearing the words, but also the expectations towards the speed made it confusing to know, when to ask for a specific advice:

> It worked well but it was... I probably thought it would go faster when I wanted say something at once like or, ask..clarification! (–) I asked once, yes 'I didn't hear can you repeat' and then..he repeated but then I didn't hear anyway (–) then he said something that it was over or something but, that..that..I didn't hear the words. (Stella)

5 Discussion

In this study, we have investigated older adults' expectations, challenges and experiences in interacting with conversational agents while cooking. We explored this specific domestic task with a three-stage research approach consisting of participatory design, a Wizard of Oz study, and qualitative interviews. Although previous studies have taken a qualitative approach in exploring older adults' interaction with and use of conversational agents [5–7, 16–19], these studies have not applied participatory design to explore their self-identified needs and perceptions and apply them to a design process. Studies utilising participatory design in the development of socially assistive robots among older adults [10, 11], on the other hand, have not focused on conversational interaction.

Our study has started from the premise that conversational agents for cognitive assistance should be designed with older adults, and should reinforce their skills and competences, encourage their participation, and strengthen their sense of security, rather than reinforce their experience of dependency [15]. Therefore, we collaboratively designed an interaction paradigm focused on collaborative assistance. Here, we answer the research questions and discuss these findings in relation to previous studies on conversational agents and older adults.

5.1 Social Connectedness with Conversational Agent

As a response to RQ1, the study shows that older adults consider conversational advice for recipe following and nutrition advice as a core functionality of a kitchen-focused agent. Tracking nutritional intake, receiving personalised recipe recommendations, advising the user on appropriate ingredients, reminding them of dietary intake, and optimising their usage of existing ingredients are cognitive tasks that older adults themselves consider appropriate and meaningful for the agent to assist with.

Conversational agents for older adults are often developed with an assumption that they can, and should, provide social connectivity and companionship in to ease social isolation [26], facilitate the receiving support from a caregiver [5] or function as a 'well-being' coach [27]. Older adults themselves, on the other hand, mostly consider these agents as tools for the purpose of information seeking [17]. Findings from all three data sets collected in this study confirm that older adults do not initially consider conversational agents as social partners, which was shown in both as explicit statements and a lack of social responses when interacting with the agent. Home cooking itself was considered a social activity between family members, but older adults did not want to have a conversation with the agent outside the immediate cooking task. This tendency to personify conversational agents or consider them as social entities seems to be more common for younger users than older ones [28, 50]. Future studies focusing on the development of conversational agents for social companionship should have a stronger emphasis on older adults' own understandings of companionship, and critically evaluate the purpose and consequences of such technologies for the experience of social connectedness.

Our study showed that older adults tend to use polite interjections when interacting with the agent, which has previously been recognised as an interaction pattern typical for older adults [25]. Polite interactions conversational agents have been connected to personifying behaviours. When users communicate with phrasings such 'please' or 'thank you', they are said to tend to anthropomorphise and attach more human-like features to the agent [6]. Our data, consisted of video recordings combined with interviews, demonstrated that although older adults communicated with polite greetings with the agent, they experienced the sociality of the agent in relation to task-based performance, rather than social connectedness as such. In some cases, polite greetings and responses from the agent facilitated user's willingness to ask advice more often, which indicates the importance of

linguistic style in designing conversa- tional agents to facilitate long-term user engagement. This points to the use of these linguistic styles as less of an indication of anthropomorphisation, and more as a recognition that entrained styles of language are persistent across human–human to human–machine interactions and that conversational mirroring—where one party aligns their conversational style with the other—works to facilitate communication in both styles of interaction.

5.2 Multimodality of Conversational Agent and Older Adults

As a response to RQ2, the study demonstrates that when interacting with the agent, older adults show challenges in confirmation and repetition, questioning and correcting, lack of conversational responses, and difficulties in hearing and understanding the multi-modal interaction. Older adults used repetition in both asking and responding to commands with the agent, and they frequently adjusted, questioned or corrected the advice received from the agent. They used rather complex queries that were all focused on prioritising task efficiency, rather than having a social interaction with the agent. Older adults encountered a frequent miscommunication with the agent, in the form of not hearing the voice commands, misinterpreting the commands or asking too many commands from the agent in a short time.

Gaze has become a popular interaction modality in studies on Human–Robot interaction to study, for instance, how gaze can facilitate task performance or social connections in small groups [50, 51]. Only few studies have investigated the gaze interaction with older adults [50, 52, 53]. Some studies show that gaze has a lower impact on task performance on older adults than middle-aged or younger participants [50, 52]. In our study, gaze interaction did not have any significant influence on the interaction strategies or the experiences of participants. Those older adults who had the gaze modality activated did not raise any positive or negative concerns during the interviews, which could indicate that the significance of gaze was relatively low for this age group.

One reason for the low impact of gaze interaction could be that gaze, as an interaction modality, is relatively unfamiliar for older adults, especially in comparison to voice interaction. Older adults may prefer voice over gaze, which was shown in the design workshop. During complex tasks such as cooking a meal from the recipe, there are several activities to focus on, and thus noticing the gaze behaviour of system may require more visual attention than the users are able or willing to give. Some studies also suggest that age-related physical changes may decrease the ability for gaze following [52]. Another studies indicate that age-related differences in gaze following are apparent in social signals but not when communicating social information [53]. The lack

of impact on task performance or they had any positive or negative experiences of it suggests that, for this age group, gaze interaction should be perhaps designed in a different way.

5.3 Age-Specificity of Interaction with Conversational Agent

As a response to RQ3, the study highlights that older adults perceived the agent to be a practical support for daily tasks, rather than a conversational partner. Older adults mentioned maintaining a clear boundary on what kind of discussion they would like to have with the agent, but considered it important for the agent to provide encouragement and social support to a certain extent. They considered the major reason for interaction failures to be caused by their lack of previous experience in using voice interaction. They also perceived a need to adjust their own interactions to be appropriate and understandable for the agent.

These findings address the importance to study the age-specificity of human–robot interaction to understand how and why older adults interact with conversational agents in a way that is considered purposeful for them. Hence, the study proposes that investigating the age-specificity in HRI should go beyond just recognising the age-related differences between certain age groups [50, 52] towards understanding how different interaction patterns and styles are connected to ageing, and why do older adults experience the interaction in a certain way in relation to their previous experience with technologies. Studies focusing on the age-specificity of human–robot interaction should also consider social and cognitive aspects of ageing, rather than being limited to physical capabilities of older adults. This has a possibility to shape the design of conversational agents to be inclusive and respectful towards older adults' skills, capabilities and participation [15].

Previous studies on conversational agents and older adults often start with the notification that older adults skills and capabilities deteriorate with age [50]. Our study, on the other hand, has highlighted older adults' own activity in shaping the functionality of conversational agents, and showing how they proactively participate as questioning or correcting the advice received from the agent for their own individual purposes. Furthermore, our study stresses the well-recognised importance of involving participants from the age group of 65 and over in technology design and development [8, 9]. Only by involving older adults—not only as test participants but active producers of interaction scripts and designs—we can ensure that these technologies eventually fit with their own understandings of appropriate technologies for conversational assistance.

5.4 Limitations

Our study has investigated older adults' interaction with a multi-modal conversational wizarded agent in a laboratory setting. While there is much to learn from observing interactions within the individual, shared and varied home environments of users on how such systems can fit with the wider socio-technical context, such studies in controlled environments allow the focus to be on the interaction and enable direct comparison between conditions and user behaviours. The use of a wizard facilitates the rapid iteration of interaction designs, enabling researchers to triangulate what development work is worthwhile and meaningful for the user group. However, the interactions challenges related directly to the implementation challenges of such systems are unable to be effectively explored or addressed in this way. Future research should focus on investigating the interaction challenges with an automated system, which would provide the opportunity to deploy it in users homes and understand how the interactions and experiences with conversational agents change over time.

5.5 Conclusions

This study has demonstrated that older adults consider conversational agents as beneficial for providing personalised advice for recipe following and nutrition advice, but voice and gaze as an interaction modality remains a challenge for this age group. Older adults' interaction with the agent was characterised by confirming and repetition, questioning and correcting, lack of conversational responses, and difficulties in hearing and understanding the multi-modal interaction. To improve the inclusiveness and accessibility of conversational agents, future research should focus on improving the mechanisms to integrate the outcomes of participatory design to technical development, and investigate older adults' interaction and experiences with an automated system in a real-life environment. The purpose and activity for which the conversational agents are designed for should promote equality and dignity in ageing. Designing conversational agents in elderly care remains one of the most ethically challenging environments [9], which cannot be solved only by focusing in technical improvement of such technologies. Participatory methods need further development in order to facilitate older adults to see themselves as designers, not only as users of technology.

Declarations

Conflict of interest The authors declare no conflict of interest and no competing interests.

Ethical Approval All participants have given an informed consent for the study, and an approval for publication of the anonymised extractions from the data. All participants were given pseudonyms, and their faces were blurred to ensure their personal integrity. The study has not collected or processed any sensitive personal data (health-related information, ethnicity, political or religious beliefs, membership in trade union, sexual orientation nor genetic information). Reference number for ethical approval: 2022-09-21.

Consent to publish All participants signed an informed consent form regarding publishing their data and photographs in an anonymised form in scientific publications.

References

1. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, Surian D, Gallego B, Magrabi F, Lau AYS, Coiera E (2018) Conversational agents in healthcare: a systematic review. J Am Med Informat Assoc 25(9):1248–1258. https://doi.org/10.1093/jamia/ocy072

2. Bentley F, Luvogt C, Silverman M, Wirasinghe R, White B, Lottridge D (2018) Understanding the long-term use of smart speaker assistants. Proc ACM Interact Mobile Wear Ubiq Technol 2(3):1–24. https://doi.org/10.1145/3264901

3. Porcheron M, Fischer JE, Reeves S, Sharples S (2018) Voice Interfaces in Everyday Life. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp. 1–12. ACM. https://doi.org/10.1145/3173574.3174214

4. Nikitina S, Callaioli S, Baez M (2018) Smart conversational agents for reminiscence. In: Proceedings of the 1st international workshop on software engineering for cognitive services, pp. 52–57. ACM (2018). https://doi.org/10.1145/3195555.3195567.

5. Zubatiy T, Vickers KL, Mathur N, Mynatt ED (2021) Empowering dyads of older adults with mild cognitive impairment and their care partners using conversational agents. In: Proceedings of the 2021 CHI conference on human factors in computing systems, pp. 1–15. ACM. https://doi.org/10.1145/3411764.3445124.

6. Pradhan A, Findlater L, Lazar A (2019) "Phantom friend" or "Just a box with information": personification and ontological categorization of smart speaker-based voice assistants by older

adults. Proceedings of the ACM on human-computer interaction 3(CSCW), 1–21 (2019). https://doi.org/10.1145/3359316.

7. Gollasch D, Weber G (2021) Age-related differences in preferences for using voice assistants. In: Mensch und computer 2021, pp. 156–167. ACM. https://doi.org/10.1145/3473856.3473889.

8. Frennert S, Aminoff H, Ostlund B (2021) Technological frames and care robots in eldercare. Int J Soc Robot 13(2):311–325. https://doi.org/10.1007/s12369-020-00641-0

9. Søraa RA, Tøndel G, Kharas MW, Serrano JA (2022) What do older adults want from social robots? A qualitative research approach to human–robot interaction (HRI) studies. Int J Soc Robot. https://doi.org/10.1007/s12369-022-00914-w

10. Lee HR, Sabanovic S, Chang W-L, Nagata S, Piatt J, Bennett C, Hakken D (2017) Steps toward participatory design of social robots: mutual learning with older adults with depression. In: Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction, pp. 244–253. ACM. https://doi.org/10.1145/2909824.3020237.

11. Sabanovi´c S, Chang W-LBennett CC, Piatt JA, Hakken D (2015) A robot of my own: participatory design of socially assistive robots for independently living older adults diagnosed with depression. In: Zhou, J., Salvendy, G. (eds.) Human aspects of IT for the aged population. design for aging vol 9193, pp. 104–114. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20892-3_11

12. Kosch T, Wozniak PW, Brady E, Schmidt A (2018) Smart kitchens for people with cognitive impairments: a qualitative study of design requirements. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp. 1–12. ACM. https://doi.org/10.1145/3173574.3173845.

13. Kosch T, Wennrich K, Topp D, Muntzinger M, Schmidt A (2019) The digital cooking coach: using visual and auditory in-situ instructions to assist cognitively impaired during cooking. In: Proceedings of the 12th ACM international conference on pervasive technologies related to assistive environments, pp. 156–163. ACM. https://doi.org/10.1145/3316782.3321524

14. Chai MZ, Soro A, Roe P, Brereton M (2017) Cooking together at a distance: sustain connectedness for long distance famiilies. In: Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems, pp. 2437–2444. ACM. https://doi.org/10.1145/3027063.3053183.

15. Kuoppam¨aki S, Tuncer S, Eriksson S, McMillan D (2021) Designing kitchen technologies for ageing in place: a video study of older adults' cooking at home. Proc ACM Interact Mobile Wear Ubiq Technol 5(2):1–19. https://doi.org/10.1145/3463516

16. Kim S, Choudhary A (2021) Exploring older adults perception and use of smart speaker-based voice assistants: a longitudinal study. Comput Hum Behavior 124:106914. https://doi.org/10.1016/j.chb.2021

17. Pradhan A, Lazar A, Findlater L (2020) Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput-Hum Interact 27(4):1–27. https://doi.org/10.1145/3373759

18. Sayago S, Neves BB, Cowan BR (2019) Voice assistants and older people: some open issues. In: Proceedings of the 1st international conference on conversational user interfaces—CUI '19, pp. 1– 3. ACM Press. https://doi.org/10.1145/3342775.3342803.

19. Li J, Qu NZ, Penaranda Valdivia K (2022) Design of a social media voice assistant for older adults. In: Cavallo F, Cabibihan J-J, Fiorini L, Sorrentino A, He H, Liu X, Matsumoto Y, Ge SS (eds.) Social robotics vol. 13817, pp. 75–88. Springer, Cham. https://doi.org/10.1007/978-3-031-24667-8_7

20. Azevedo RFL, Morrow D, Graumlich J, Willemsen-Dunlap A, Hasegawa-Johnson M, Huang TS, Gu K, Bhat S, Sakakini T, Sadauskas V, Halpin DJ (2018) Using conversational agents to explain medication instructions to older adults. In: AMIA annual symposium proceedings. AMIA Symposium 2018, 185–194

21. Bennion MR, Hardy GE, Moore RK, Kellett S, Millings A (2020) Usability, acceptability, and effectiveness of web-based conversational agents to facilitate problem solving in older adults: controlled study. J Med Internet Res 22(5):16794. https://doi.org/10.2196/16794

22. Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, Briatore A, Coiera E (2019) The personalization of conversational agents in health care: systematic review. J Med Internet Res 21(11):15360. https://doi.org/10.2196/15360

23. Morrow DG, Lane HC, Rogers WA (2021) A framework for design of conversational agents to support health self-care for older adults. Hum Factors J Human Factors Ergon Soc 63(3):369–378. https://doi.org/10.1177/0018720820964085

24. Blair J, Abdullah S (2019) Understanding the needs and challenges of using conversational agents for deaf older adults. In: Conference companion publication of the 2019 on computer supported cooperative work and social computing, pp. 161–165. ACM. https://doi.org/10.1145/3311957.3359487

25. Chung K, Oh YH, Ju DY (2019) Elderly users' interaction with conversational agent. In: Proceedings of the 7th international conference on human-agent interaction, pp. 277–279. ACM (2019). https://doi.org/10.1145/3349537.3352791

26. Simpson J, Gaiser F, Mack M, Bregott T (2020) Daisy: a friendly conversational agent for older adults. In: Proceedings of the 2nd conference on conversational user interfaces, pp. 1– 3. ACM (2020). https://doi.org/10.1145/3405755.3406166

27. El Kamali M, Angelini L, Caon M, Khaled OA, Mugellini E, Dulack N, Chamberlin P, Craig C, Andreoni G (2020) NESTORE: Mobile chatbot and tangible vocal assistant to support older adults' well- being. In: Proceedings of the 2nd conference on conversational user interfaces, pp. 1–3. ACM (2020). https://doi.org/10.1145/3405755.3406167

28. Druga S, Williams R, Breazeal C, Resnick M (2017) "Hey Google is it OK if I eat you?": Initial explorations in child-agent interaction. In: Proceedings of the 2017 conference on interaction design and children, pp. 595–600. ACM (2017). https://doi.org/10.1145/3078072.3084330

29. Abdolrahmani A, Kuber R, Branham SM (2018) "Siri talks at you": An Empirical investigation of voice-activated personal assistant (VAPA) usage by individuals who are blind. In: Proceedings of the 20th international ACM SIGACCESS conference on computers and accessibility, pp. 249–258. ACM (2018). https://doi.org/10.1145/3234695.3236344.

30. Hamada R, Okabe J, Ide I, Satoh S, Sakai S, Tanaka H (2005) Cooking navi: assistant for daily cooking in kitchen. In: Proceedings of the 13th Annual ACM International conference on multimedia—MULTIMEDIA '05, p. 371. ACM Press. https://doi.org/10.1145/1101149.1101228

31. Doman K, Kuai CY, Takahashi T, Ide I, Murase H (2012) Smart VideoCooKing: a multimedia cooking recipe browsing application on portable devices. In: Proceedings of the 20th ACM international conference on multimedia—MM '12, p. 1267. ACM Press. https://doi.org/10.1145/2393347.2396435.

32. Sato, A., Watanabe, K., Rekimoto, J.: MimiCook: a cooking assistant system with situated guidance. In: Proceedings of the 8th international conference on tangible, embedded and embodied interaction—TEI '14, pp. 121–124. ACM Press (2013). https://doi.org/10.1145/2540930.2540952.

33. Frummet A, Elsweiler D, Ludwig B (2022) What can i cook with these ingredients?: Understanding cooking-related information needs in conversational search. ACM Trans Inf Syst 40(4):1–32. https://doi.org/10.1145/3498330

34. Lim H, Huang X, Miller S, Edelmann J, Euken T, Voida S (2019) Smart cook: making cooking easier with multimodal learning. In: Adjunct Proceedings of the 2019 ACM international joint conference on pervasive and ubiquitous computing and proceedings of

the 2019 ACM international symposium on wearable computers, pp. 129–132. ACM. https://doi.org/10.1145/3341162.3343836

35. Chang M, Truong A, Wang O, Agrawala M, Kim J (2019) How to design voice based navigation for how-to videos. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp. 1–11. ACM. https://doi.org/10.1145/3290605.3300931

36. Chang M, Huh M, Kim J (2021) RubySlippers: supporting content-based voice navigation for how-to videos. In: Proceedings of the 2021 CHI conference on human factors in computing systems, pp. 1–14. ACM. https://doi.org/10.1145/3411764.3445131

37. Randall N, Sabanovic S, Chang W (2018) Engaging older adults with depression as co-designers of assistive in-home robots. In: Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare, pp. 304–309. ACM. https://doi.org/10.1145/3240925.3240946

38. Bajones M, Fischinger D, Weiss A, Puente PDL, Wolf D, Vincze M, Krtner T, Weninger M, Papoutsakis K, Michel D, Qammaz A, Panteleris P, Foukarakis M, Adami I, Ioannidi D, Leonidis A, Antona M, Argyros A, Mayer P, Panek P, Eftring H, Frennert S (2020) Results of field trials with a mobile service robot for older adults in 16 private households. ACM Trans Human-Robot Interact 9(2):1–27. https://doi.org/10.1145/3368554

39. Gasteiger N, Ahn HS, Gasteiger C, Lee C, Lim J, Fok C, Macdonald BA, Kim GH, Broadbent E (2021) Robot-delivered cognitive stimulation games for older adults: usability and acceptability evaluation. ACM Trans Human-Robot Interact 10(4):1–18. https://doi.org/10.1145/3451882

40. Pou-Prom C, Raimondo S, Rudzicz F (2020) A conversational robot for older adults with Alzheimers disease. ACM Trans Human-Robot Interact 9(3):1–25. https://doi.org/10.1145/3380785

41. Jung MM, Ludden GDS (2019) What do older adults and clinicians think about traditional mobility aids and exoskeleton technology? ACM Trans Human-Robot Interact 8(2):1–17. https://doi.org/10.1145/3311789

42. Palumbo V, Patern F (2020) Serious games to cognitively stimulate older adults: a systematic literature review. In: Proceedings of the 13th ACM International conference on pervasive technologies related to assistive environments, pp. 1–10. ACM. https://doi.org/10.1145/3389189.3393739.

43. McMillan D, Brown B, Kawaguchi I, Jaber R, Solsona Belenguer J, Kuzuoka H (2019) Designing with Gaze: Tama—a Gaze Activated Smart-Speaker. Proc ACM Hum-Comput Interact 3:1–26. https://doi.org/10.1145/3359278

44. Heerink M, Krse B, Evers V, Wielinga B (2010) Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J Soc Robot 2(4):361–375. https://doi.org/10.1007/s12369-010-0068-5

45. Borsci S, Malizia A, Schmettow M, van der Velde F, Tariverdiyeva G, Balaji D, Chamberlain A (2022) The chatbot usability scale: the design and pilot of a usability scale for interaction with AI-based conversational agents. Pers Ubiquit Comput 26(1):95–119. https://doi.org/10.1007/s00779-021-01582-9

46. Liu Y-C, Chen C-H, Lin Y-S, Chen H-Y, Irianti D, Jen T-N, Yeh J-Y, Chiu SY-H (2020) Design and usability evaluation of mobile voice-added food reporting for elderly people: randomized controlled trial. JMIR mHealth uHealth 8(9):20317. https://doi.org/10.2196/20317

47. Vaismoradi M, Turunen H, Bondas T (2013) Content analysis and the-matic analysis: implications for conducting a qualitative descriptive study: qualitative descriptive study. Nurs Health Sci 15(3):398–405. https://doi.org/10.1111/nhs.12048

48. Clark L, Pantidi N, Cooney O, Doyle P, Garaialde D, Edwards J, Spillane B, Gilmartin E, Murad C, Munteanu C, Wade V, Cowan BR (2019) What makes a good conversation?: Challenges in designing truly conversational agents. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp. 1–12. ACM. https://doi.org/10.1145/3290605.3300705

49. Rudzicz F, Wang R, Begum M, Mihailidis A (2015) Speech interaction with personal assistive robots supporting aging at home for individuals with alzheimers disease. ACM Trans Accessible Comput 7(2):1–22. https://doi.org/10.1145/2744206

50. Seaborn K, Sekiguchi T, Tokunaga S, Miyake NP, Otake-Matsuura M (2022) Voice over body? Older adults' reactions to robot and voice assistant facilitators of group conversation. Int J Soc Robot. https://doi.org/10.1007/s12369-022-00925-7

51. Tuncer S, Gillet S, Leite I (2022) Robot-mediated inclusive processes in groups of children: from gaze aversion to mutual smiling gaze. Front Robot AI 9:729146. https://doi.org/10.3389/frobt.2022.729146

52. Morillo-Mendez L, Schrooten MGS, Loutfi A, Mozos OM (2022) Age-related differences in the perception of robotic referential gaze in human-robot interaction. Int J Soc Robot. https://doi.org/10.1007/s12369-022-00926-6

53. Pavic K, Oker A, Chetouani M, Chaby L (2021) Age-related changes in gaze behaviour during social interaction: an eye-tracking study with an embodied conversational agent. Quart J Exp Psychol 74(6):1128–1139. https://doi.org/10.1177/1747021820982165

Sanna Kuoppamäki (PhD) is an Assistant Professor in Technology in Health Care at the Department of Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, Sweden. Her research investigates the design and use of interactive technologies for health and well-being from the age and life course perspective.

Razan Jaberibraheem (MSc) is a PhD student at Department of Computer and System Sciences, Stockholm University, Sweden. Her research interests include conversational user interfaces, speech and gaze interaction modalities and machine learning in the field of human-computer interaction.

Mikaela Hellstrand (Msc) is a PhD student at the Department of Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, Sweden. Her research applies ethnographic methods to investigate the implementation of data-driven services in elder care facilities from the perspective of patients and their caregivers.

Donald McMillan (PhD) is a senior lecturer at the Department of Computer and System Sciences, Stockholm University, Sweden. His research develops human-computer and human-robot interaction paradigms through the application of machine learning.