Computers in Human Behavior 124 (2021) 106914

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study

Sunyoung Kim *, Abhishek Choudhury

School of Communication and Information, Rutgers University, New Jersey, USA

ARTICLE INFO

Keywords:
Older adults
Voice assistant
Smart speaker
Longitudinal study

ABSTRACT

Thanks to their conversational capabilities, smart speaker-based voice assistants are gaining attention for their potential to support the aging population, though the empirical evidence is still scarce. This paper aims to obtain empirical evidence on older adults' experiences with a voice assistant. We especially focused on how their perception and use change over time as they progress from novice to more experienced users through a longitudinal field deployment study. We deployed Google Home devices in the homes of twelve older adults aged 65 and above and studied their use for sixteen weeks. Results show that the benefits our participants perceived have incrementally changed from enjoying simplicity and convenience of operation in the early phase of the study to not worrying about making mistakes and building digital companionship as they got used to using it. Results also show that participants confronted several challenges that evolved from the unfamiliarity with a voice assistant in their first interactions to coping with the functional errors due to limited speech technology as they got used to using it. Based on the results, we discuss design implications that could foster better user experiences with a voice assistant among older adults.

1. Introduction

With the advancement of artificial intelligence and speech technology, smart speaker-based voice assistants are becoming increasingly available in the market. It is estimated that in 2019, over 98 million units of smart speakers have been sold worldwide, and it is predicted to reach up to 409.4 million units in 2025 (Vailshery, 2019). Nearly a quarter of households own a smart speaker in the US, and more than half of them use two or more smart speakers (Richter, 2020). The smart speaker is a type of speaker (e.g., Amazon Echo, Google Home) with an integrated virtual assistant (e.g., Amazon Alexa, Google assistant) that responds to voice commands. It assists users in their daily lives, such as playing music, checking weather forecasts, setting alarms and reminders, controlling applicable smart home devices, and answering general questions. Because speech is one of the most natural ways of human communication, using speech to interact with devices can lower the barriers of technology use for those who are less familiar or have manual-dexterity and vision-related issues with typing- and screen-based interfaces. Therefore, this technology has gained particular attention as beneficial for older adults (Blair & Abdullah, 2019; Portet, Vacher, Golanski, Roux, & Meillon, 2013). Recent studies have shown that older adults generally have positive perceptions when introduced to a smart speaker (Blair & Abdullah, 2019), preferring voice-based user interfaces over traditional interaction modalities such as clicking or typing (Kowalski et al., 2019; Wulf et al., 2014).

Despite the rapid growth of voice technology and its hyped anticipation for older adults, smart speakers' actual adoption and use among older adults are very low. Older adults are slower to adopt new technologies than younger adults (Vaportzis, Giatsi Clausn, & Gow, 2017), and a smart speaker is not an exception. The statistics show that younger Americans in the 18–29 age group are 75% more likely to own a smart speaker than those over 60 in 2019 (Kinsella, 2019). Considering the potential of voice assistants for older adults and their low adoption rate, it is crucial to investigate the use of voice assistants for older adults.

To date, researchers have extensively investigated various aspects of using voice assistants, including the use by specific user groups, (e.g., people with disabilities (Abdolrahmani, Kuber, & Branham, 2018; Pradhan et al., 2018), children (Druga et al., 2017; Garg & Sengupta, 2020), low-income populations (Robinson et al., 2018), privacy concerns (Lau et al., 2018; Liao et al., 2019), conversational aspects (Purington et al., 2017; Vtyurina & Fourney, 2018), and personification (Lopatovska & Williams, 2018; Pradhan et al., 2019). Recently,

* Corresponding author.
E-mail addresses: Sunyoung.kim@rutgers.edu (S. Kim), ac949@scarletmail.rutgers.edu (A. Choudhury).

https://doi.org/10.1016/j.chb.2021.106914
Received 15 January 2021; Received in revised form 2 June 2021; Accepted 9 June 2021
Available online 14 June 2021

researchers have recognized the potentials that a smart speaker can offer in the aging society (Portet, Vacher, Golanski, Roux, & Meillon, 2013) and investigated older adults' experiences with this technology. For instance, Pradhan et al. found reliability concerns as one of the key challenges that prevent older adults with low technology use from adopting this technology (Pradhan et al., 2020), and Trajova and Martin-Hammond found that lack of perceived utility contributes to increasing attrition rates among experienced older users (Trajkova & Martin-Hammond, 2020). Yet, missing is a holistic perspective on how older adults progress from novice to more experienced users and what contributes to or prevents their sustained use of a voice assistant. This paper contributes to the growing body of literature by investigating how novice older adults' perception and use of a voice assistant change over time as they become more experienced through a longitudinal deployment study. Specifically, we aim to answer the following research questions:

- What do older adults use a voice assistant for?
- What different benefits do they perceive as they use the device over time?
- What different challenges do they face when using the device over time, and how do they progressively respond to or cope with those challenges?

To answer our research questions, we deployed a smart speaker in the homes of twelve older adults aged 65 and above and studied its use for sixteen weeks. We conducted in-person interviews by visiting the homes every other week for the deployment duration and collected usage logs from the system's history repository for data collection.

Our findings revealed the frequently used features, perceived and experienced benefits, and difficulties with coping strategies that older adults might have when interacting with a voice assistant. We found that participants recognized, appreciated, and enjoyed several benefits that a voice assistant offers throughout the study period. However, they did not acknowledge all the benefits in their first interactions. Instead, the perceived benefits have incrementally changed *from enjoying simplicity and convenience of operation* in the early phase of the study to *not worrying about making mistakes* and *building digital companionship* as they got used to using it. In addition, we found that the driving force to sustain the use was how positive the user experience with a voice assistant was, which was determined not only by its task completion but also through building companionship with the device. Then, the positive experience with a voice assistant contributed to developing a resilient attitude toward its functional errors. These findings can be used as design guidelines to better leverage and promote the sustained use of the emerging personal technology, a smart speaker–based voice assistant, to support the aging society. To the best of our knowledge, this is the first study that investigated the progressive use of a voice assistant among older adults as they move from novice to more experienced users through a longitudinal field deployment study.

2. Background

The technology that responds to voice commands has been called by many different terms, including voice assistant, voice-user interface, virtual assistant, intelligent assistant, and conversational agent. This paper uses "voice assistant" to encompass the above terminology and refer to a smart speaker's voice-based interface.

2.1. Voice assistants

In recent years, there has been growing interest in smart speaker-based voice assistants in the Human-Computer Interaction community, studying the use of voice assistants from several angles. First, the lack of a detailed understanding of how these technologies are used due to their novelty has led to investigating general patterns of using voice assistants in everyday contexts through a quantitative analysis of voice command data (Sciuto et al., 2018). As a result, researchers have identified music, search, and Internet of Things (IoT) as the most frequently used features of voice-based commands (Ammari et al., 2019), popular types of commands in different times of the day (e.g., entertainment and home automation commands peak in the evening while the weather and time request peak in the early morning hours) (Bentley et al., 2018), and voice assistants integrating into people's everyday routines by embedding in a range of conversational activities in the home (e.g., family dinners) (Porcheron, Fischer, Reeves, & Sharples, 2018). Other researchers have found that individuals are motivated to use a voice assistant by utilitarian, symbolic, and social benefits (McLean & Osei-Frimpong, 2019). However, its attrition rate is high due to unmet expectations and a lack of perceived utility (Cho et al., 2019).

Privacy was another major research topic due to a voice assistant's privacy-intrusive potential to continuously listen to voices in intimate spaces such as the home (Clark et al., 2019; Cowan et al., 2017; Myers et al., 2018). The conversational nature and the "always-on" listening feature of a smart speaker have also drawn researchers' attention to social aspects of using voice assistants, such as personification or anthropomorphism (Pfeuffer, Alexander, Gimpel, & Hinz, 2019). Researchers have revealed that some users perceived the device as a companion or a friend and exhibited personifying behaviors toward voice assistants by using human-like terms such as "she" or "her" when referring to a voice assistant (Purington et al., 2017). At the same time, others have argued that such interaction patterns are "mindless" social responses that may not relate to the actual perception of personification (Lopatovska & Williams, 2018). Relatedly, researchers have investigated factors that constitute an effective conversation with voice assistants, though user experiences with voice assistants thus far remained disappointing due to the constrained and predefined turn-taking structure of question-answer rather than a realistic dialog of conversation (Cowan et al., 2017; Murad & Munteanu, 2019).

Lastly, research has explored the potential utility of voice assistants as assistive technology for people with disabilities (Abdolrahmani, Kuber, & Branham, 2018), specific health concerns (Maharjan et al., 2019), and the aging population. Recently, research efforts have been increasingly devoted to understanding the use of smart speaker-based voice assistants by and for older adults, which we describe in the next section.

2.2. Voice assistants for older adults

An extensive research effort has been made to identify older adults' needs and evaluate the feasibility, effectiveness, and acceptability of existing technologies to meet their needs (Pyae & Joelsson, 2018; Liu et al., 2016; Mitzner et al., 2010). Many studies have reported that older adults are generally positive about new technologies and are willing to accept them if perceived benefits are evident (Morris et al., 2013). However, a digital divide still exists (Delello & Mcwhorter, 2017), and more efforts are needed to make new technologies readily accessible to older adults.

Research has shown that voice-based interactions have several potential benefits to support older adults. Common difficulties older users experience when using computers and smartphones are caused by the prevalence of Graphical User Interface (GUI) and its desktop metaphor (Sayago et al., 2019). Using voice as an interaction modality can help address many difficulties that GUI entails. First, through speech, voice assistants allow people to interact with them in a natural way of communicating with a person (Mctear et al., 2016). Because they are deemed simple and easy to use and ideal for users with visual and motor impairment, speech has been considered an accessible and useable interaction modality for older adults (Bickmore et al., 2005). Yet, several challenges also exist, such as unfamiliarity with talking to a device (Myers et al., 2018) and not knowing what (and how) to say to it (Sayago et al., 2019), as well as aging-related declines that might impede using

voice assistants. Hearing loss, which is one of the most common physical problems that people experience as they age, imposes a fundamental challenge for its use. Designers are usually considerably younger and may not know about the physical and psychological aspects of aging, having grown up using more advanced technologies than older adults (Porcheron, Fischer, Reeves, & Sharples, 2018; Sebastiaan et al., 2016). Thus, it is essential to investigate the barriers that older adults might have when using speech to interact with a device from the perspectives of older adults.

In the context of voice technology for aging, most studies have focused on the idea of facilitating voice-based systems as digital companions or virtual assistants to support aging in place (Bickmore et al., 2016; Heerink et al., 2010; Tsiourti et al., 2016; Van Hoof, Kort, Rutten, & Duijnstee, 2011). In particular, a stream of research has investigated the benefits of using embodied conversational agents – a graphical human agent capable of engaging in conversations with humans by both understanding and producing speech and facial expressions (Cassell, Bickmore, Campbell, & Vilhjálmsson, 2000, pp. 29–63) – to support older adults' healthcare in various aspects (e.g., improving access to online health information (Bickmore et al., 2016), delivering medication instructions (Azevedo et al., 2018), mitigating social isolation (Sidner et al., 2018)). Recently, researchers have recognized the potentials that a smart speaker can offer in the aging society (Pradhan et al., 2020) and investigated older adults' use of this technology.

In summary, older adults were positive about the usability of a voice assistant when they were first introduced to it, thanks to its simple and effortless interactions (Kowalski et al., 2019; Portet, Vacher, Golanski, Roux, & Meillon, 2013). At the same time, they had raised many open questions about its usefulness and utility as they were using it (Pradhan et al., 2020; Trajkova & Martin-Hammond, 2020). Yet, missing is a holistic perspective on how older adults' perception and experience change over time and what contributes to or prevents their sustained use of a voice assistant. This paper contributes to this emerging body of literature by exploring how older adults perceive and use a voice assistant progressively as they move from novice to more experienced users through a longitudinal field deployment study.

3. Methods

Our analysis is based on the data collected from the deployment of Google Home devices in the homes of twelve individuals aged 65 or older up to sixteen weeks. The study was approved by local Institutional Review Boards, and informed consent was obtained from all participants for the use of collected data before participating in the study.

3.1. Participants

For participant recruitment, we first contacted two senior-living communities located in the greater New York area. We visited each community and explained the purpose of the study to the manager. Upon their approval, we posted a recruitment flyer in the lobby of the communities. Two inclusion criteria for participation were age being over 65 and having no prior experience with a voice assistant.

In total, we recruited 12 participants (7 females and 5 males), ranging in age from 65 to 95 (mean age = 83.8, SD = 9.1, see Table 1). All participants resided in a single-person unit in the community. About general technology use, all participants said that they were familiar with computers, tablets, and a smartphone, and four participants said that they have seen a smart speaker in their children's homes but have not used it. Seven participants owned a tablet, and all participants reported regularly using a computer for information search and email. About health conditions, two participants were wearing a hearing aid but did not have any problem with having a conversation. Three participants were using a wheelchair due to various joint issues. Other than those residents, all participants reported being healthy both physically and cognitively with mild aging-related health issues, such as arthritis or mild memory loss. Among them, three participants dropped out of the study within the first month of deployment due to personal reasons. One participant dropped after having three interviews because she was hospitalized, and two participants dropped after having two and three interviews respectively because they lost interest in using a voice assistant. The other nine participants completed the study for the entire sixteen weeks.

Table 1
Participant demographics.

ID	Age	Gender	Health concerns	Length of participation	Reason for dropping
P1	90	M	Arthritis (on a wheelchair)	Full	–
P2	77	F	–	Full	–
P3	65	F	Mild memory loss	Full	–
P4	75	M	Mild memory loss	Full	–
P5	94	F	Wearing a hearing aid	Full	–
P6	87	F	Wearing a hearing aid	Full	–
P7	89	M	Mild memory loss	Full	–
P8	78	F	–	Full	–
P9	78	M	–	Full	–
P10	95	F	Arthritis (on a wheelchair)	4 weeks	Hospitalization
P11	85	F	–	2 weeks	Lost interest
P12	82	M	Arthritis (on a wheelchair)	4 weeks	Lost interest

3.2. Data collection

Our interview protocol focused on investigating how older adults initially perceive and respond to a voice assistant, how they use it in their daily lives, what challenges and difficulties they face when using it and how they cope with those challenges, and how their perspectives and usage patterns change over time. We constructed a set of open-ended interview questions in three phases of the study duration to explore these spaces. In the first phase, we focused on understanding initial impressions and the perceived usefulness of a voice assistant in the first few weeks after installation. In the second phase, we focused on exploring the user experience in-depth, including usage patterns, the needs and challenges, and strategies to cope with breakdowns when interacting with a voice assistant throughout the deployment duration except for the last interview. When needed, participants were asked to interact with a voice assistant during the interview to demonstrate how they would use it in their everyday lives. For instance, when a participant reported a complaint of a voice assistant not understanding their commands, we asked them to show how they would make a command during the interview to understand the context behind user feedback. In the third phase, we focused on the overall reflection on users' experiences of interacting with a voice assistant and suggestions for improvements in the last interview. While we had three phases for interview protocols, neither were these mutually exclusive nor had a fixed duration of phases in the interviews. Instead, we proceeded with the interviews flexibly depending on participants' experiences. For instance, those who quickly adapted to using a voice assistant moved to the second phase within the first couple of interviews, while those who needed more time to be familiar with using it stayed longer in the first phase.

For the first interview, the research team visited a participant's home and set up a Google Home mini in the location of their preference (e.g., a nightstand, a coffee table). After the device setup, a participant was introduced to Google Home mini as "a device that responds to your voice command, providing you answers about things you ask or need on an

everyday basis" and given basic instruction on how to use it for about 10 min. For the instruction, a researcher first demonstrated how to use the device by making voice commands for basic tasks, such as setting alarms and reminders (e.g., "Remind me to take my medicine at 10 a.m."), streaming music and radio (e.g., "Play music"), and asking general questions (e.g., "Who is the second president of the United States?"). Then, participants were prompted to try interacting with the device and to ask questions about using it. Once they had no more questions about using a voice assistant, we started to ask for feedback about their initial perspectives on the device for the rest of the interview. In addition, participants filled out a survey to inform us about their basic demographic information, including age, health concerns, and experience with technology. Lastly, participants were told to freely interact with a voice assistant as much or as little as they wanted throughout the study period and were given the contact information of the research team if they needed technical support, and we completed the first interview that lasted about an hour. Next, we conducted eight follow-up interviews by visiting them every other week for sixteen weeks, which makes nine interviews per participant. Each follow-up interview lasted between 30 min and 1.5 hours. All interviews were audio-recorded and transcribed.

In the last interview, participants were offered an option to keep a Google Home mini if they wanted, and all participants decided to keep it. Participants who completed the study were fully compensated with a $160 gift card upon completion. Those who withdrew were partially compensated with a gift card which amount was prorated by the duration of participation.

We also collected the device usage logs from Google's activity history repository, My Activity,[1] which stores the complete history of a pair of user's commands to Google's products and the response it has given. We downloaded a pair of voice inquiries made to a Google mini and its verbal responses of all participants for the duration of the study from this repository. This usage log provided us with the actual interaction patterns to complement participants' perceptions of these interactions.

3.3. Data analysis

We analyzed our interview data using thematic analysis to reveal patterns across data sets (Braun & Clarke, 2006). Thematic analysis is a method for identifying patterns and themes in qualitative data, making it suitable for analyzing our data set, where the blending of interview scripts, field notes, and photos creates a well of potentially rich thematic data to draw from. We selected thematic analysis because of its emphasis on proceeding with an open mind to investigate explanatory conceptual themes associated with older adults' use of a voice assistant over time. The thematic analysis involves open coding, axial coding, and selective coding for theme identification. To this end, each interview script was analyzed with open coding to note the themes or factors that emerged in the data. Then, the emerged themes were continuously discussed as a group with another author until no new information was anticipated. First, we conducted open coding to identify concepts significant in the data as abstract representations of events, objects, happenings, actions, etc. The example below expresses one participant's lack of confidence about interacting with a voice assistant when first introduced. This response is coded as "No self-efficacy".

[No self-efficacy] "I'd like to have one (a voice assistant), but you think I would be able to use it?" [/No self-efficacy](P2W1[2]).

Next, we categorized the related concepts created by open coding into conceptual phenomena using axial coding. Phenomena refer to repeated patterns of events, happenings, actions, and interactions representing people's responses to problems and situations. For instance, "First impression" represents a participant's reaction or response to a voice assistant when they first interacted with it. The open code "No self-efficacy" in the example above was categorized as "First impression" during axial coding. Lastly, we followed the selective coding process to assemble our conceptual phenomena extracted from axial coding. The goal of this step is to integrate all concepts through building relationships across phenomena.

In addition, the first author manually coded the usage logs deductively by the type of operation, such as playing music, asking general questions, having a casual conversation, checking time/weather, setting up a reminder/alarm, operating basic controls (e.g., "stop" to stop playing music), etc. For instance, we coded the query "play a song from the 80s" as "playing music." Since we used one Google account created by the research team for all devices, the usage log data was only available for the sample in aggregate, not per participant.

4. Findings

4.1. The features in use

In total, we retrieved 2242 pairs of request-response communications from the usage log of the four-month deployment study. Among those, 1488 pairs of communications (66.4%) were successful as a voice assistant performed tasks as requested, while 754 pairs (33.6%) were not as their responses were logged as "Sorry, I don't understand." While the mean frequency of use is 1.8 times per day, the actual usage rates greatly varied among participants. Some participants said that they barely used the device, and others used it a lot when we asked about their experience with it in the interviews. The most frequently used feature of a voice assistant was playing music (e.g., "Play sounds of the ocean," 37%), followed by searching for general information (e.g., "Is Friends playing on TBS tonight?", 16.5%), making casual conversations (e.g., "How are you?", 12.1%), checking the current time and date (e.g., "What time is it?", 11.1%), setting up a reminder or an alarm (e.g., "Remind me to take my med at 9 a.m.", 4.3%), and checking weather forecasts (e.g., "What's the weather today like?", 3.2%) (See Table 2). This echoes the findings by Ammari et al. that identified music and search as the most commonly used features of voice assistants (Ammari et al., 2019).

Playing music before bed to fall asleep and waking up checking the time and weather were the two most prevalent features of use among all participants, which has quickly become part of their everyday routines.

> "I play it (a voice assistant) before I go to bed at night, the sounds of the ocean. That helps me fall asleep. That's technology. If you don't like a song they're playing, you say, "Hey Google, I don't like that song. Play something else". And within an instant, they're playing something else. That's the part that amazes me. How can they do it that fast? They're not searching for anything. It just comes." (P5W2)

> "When I wake up, and I want to know the time I say Hey Google, what time is it. Or hey Google, how's the weather gonna be today, and she says the weather today is so and so. I used to always turn on the TV as a first thing in the morning to see the weather. I don't do that anymore, which I think is very fabulous in itself." (P8W5)

These patterns were also reflected in the frequency of operation by the time of day when the usage surged around 9 p.m. and 7 a.m. (See Graph 1). The usage also surged during the daytime between 1 p.m. and 4 p.m., but we do not have concrete evidence to explain the rise of the usage frequency in the afternoon since there was no specific pattern of use in the commands. We only assume that it might be because participants had more free time staying in their rooms between lunch and dinner while attending social activities before lunch and watching TV after dinner.

One thing to note is that the actual frequency of using a reminder or alarm must be higher than what was shown in the usage log. The log data records user activities as a pair of a user's verbal command and the

[1] https://myactivity.google.com/myactivity.

[2] In the excerpt, P# refers to the #th participant, and W# refers to the #th interview. For instance, P2W1 is an excerpt from Participant 2 in the first interview. And VA refers to a voice assistant.

Table 2
The frequency of operation by topics: Successful operations refer to the commands with proper device responses.

Operation	Music	Search	Basic device control	Casual conversation	Time	Reminder	Weather	Other	Total
All	829 (37%)	371 (16.5%)	315 (14%)	272 (12.1%)	248 (11.1%)	96 (4.3%)	72 (3.2%)	39 (1.8%)	2242
Successful	610 (41%)	207 (13.9%)	186 (12.5%)	146 (9.8%)	189 (12.7%)	82 (5.5%)	45 (3.0%)	23 (1.5%)	1488

device's voice response. And the response is blank when a user command is "stop" since there is no voice response to "stop," the command that a user makes to turn off a reminder notification or stop playing music. For recurrent reminders and alarms, a command for initial setup was captured in the log data (e.g., "Wake me up at 7 a.m. every morning."), while its repeated use throughout the study period was indirectly logged as the "stop" command. This is reflected in the frequency of commands for basic device control (14%), 52% of which was "stop" made in the morning, though we cannot attribute all these commands to using a reminder or alarm feature since it is also used to stop playing music.

4.2. The perceived and experienced benefits

Older people are a heterogeneous group, and each person has their perceptions and experiences about technology. Therefore, our participants expressed divergent feelings and perspectives about a voice assistant when first introduced, from excitement and curiosity to hesitance, uncertainty, and even refusal of the device.

> "I'd like to have one (a voice assistant), but you think I would be able to use it?" (P2W1)

> "I'm really amazed by it (a voice assistant), really and truly. I've seen a lot of things, but I think this is about the smartest. People are now living in this modern technology, and it's just part of life for them. As being older, I didn't have any of this as a kid. Here I am, at this age, at 95, and I'm listening to Google Home. It really is an accomplishment to listen to what this little thing can say. It's amazing." (P5W1)

> "Could people get any lazier? How lazy have people gotten today that they can't even stand up and turn a light on and off? That's a little sad. There's a movie about the future, and all the people are so huge that they go around on these conveyor belts, in chairs, because they're too lazy to walk from one place to the other. That's where we're headed with a technology like this. It's a little scary." (P12W1)

Then, participants recognized, appreciated, and enjoyed several benefits that a voice assistant offers as they continued interacting with it. However, they did not acknowledge all of these benefits in their first interaction. Instead, the types of benefits that participants mentioned have incrementally changed from *enjoying simplicity and convenience of operation* in the early phase of the study to *not worrying about making mistakes* and *building digital companionship* as they got used to using it. While these benefits are not surprising or novel, our findings demonstrate how older adults gradually become aware of and experience some of the benefits that a voice assistant has to offer as they progress from novice to more experienced users over time. In what follows, we report how participants' use and perceptions changed over time.

4.2.1. In the early phase among novice users: simplicity and ease of use

During the first interaction with a voice assistant, most participants instantly recognized and appreciated the simplicity and ease of using a voice assistant by its voice-based interaction modality. Several participants commented that they did not have to learn any new skills or receive training but simply had to talk to the device as an immediate and tangible benefit. A prior study showed that the perceived effort to learn a new technology is one of the critical barriers that prevent older adults from adopting it (Kim et al., 2016). For instance, having basic typing skills and understanding the meanings of GUI metaphors are necessary when using typing- and screen-based devices, which some older adults find demanding to acquire (Mitzner et al., 2010). A voice assistant successfully removed this perceived barrier and made our participants, first-time users, readily accept the device with no concerns about learning new things.

> "Computers came along, but you have to learn how to use different icons and get in and out of Windows and load things and download things. And this (a voice assistant) is like having a person around you that talks back if you want it to. It's awesome. I think I can use it." (P9W1)

> "It's cooler. I don't have to type into it, which I am not good at. I just say, hey, give me some Louis Prima or Frank Sinatra, and I get it right away. I don't have to wait. That's like so futuristic that I didn't think I'd live to see this day." (P6W2)

4.2.2. Transitioning to competent users: convenience of operating without physical interaction

After using a voice assistant for a few weeks, participants started to mention the convenience of interacting with it from a distance without having to use any tools for input or read visual outputs. Participants appreciated that a voice assistant does not require any physical interaction except speaking and listening to it. Particularly for those who had physical declines, such as decreased mobility or vision loss typical of the natural aging process, being able to operate the device using a voice from a distance was acknowledged as a significant advantage over other devices they own.

> "I like it (a voice assistant) because I can't see without my glasses. I think it is kind of in the back of my mind because I don't physically have to look at it. I'll go by and say Hey Google, what's so and so. And it tells me so fast that I don't even look at it." (P5W4)

However, such benefits entailed concerns later on. For instance, executing a command without any physical effort made them worry about becoming lazy and inactive, which would ultimately negatively impact the quality of life.

> "I love to hear it (a voice assistant) read me a book. But see, then, I'm not using my brain. I don't wanna become a vegetable and have to depend on it for everything. And I love to read. I love reading a good book. Otherwise, your brain will turn into Jell-O. You gotta use your brain." (P2W6)

> "People are just gonna get really fat because they never have to move anymore. You don't even have to stand up to make a phone call or go get a phone. I don't know if it's great. I mean it's fun and it's good, but I don't know how good that is." (P11W7)

4.2.3. Transitioning to experienced users: No worries about making mistakes

As participants became used to using a voice assistant in the later phase of the study, they increasingly expressed their emotional relief from a concern of making mistakes that they had when using personal computing devices, such as a computer or a smartphone. One of the psychological barriers that prevent older adults from adopting new technology is the fear of making operational mistakes. Prior work showed that older people tend to reject new technology because they are afraid of making mistakes, such as clicking a wrong button or deleting an

important file, which might be irreversible for themselves (Knowles & Hanson, 2018). Using speech as an interaction modality mitigated this concern as the worst possible consequence of making a wrong voice command is nothing other than the command not being executed. Thus, some participants even further experimented with error-prone commands without any concern about the consequence.

> "I don't have to worry about pressing a wrong button and delete everything. If that happens, I will have no idea how to handle it." (P1W5)

> "I like the fact that you can just ask this silly little thing no matter what it is. And that thing can tell you no matter what I ask. I don't have to worry about what if I do something wrong. So, I ask her silly questions all the time." (P8W6)

4.2.4. Among experienced users: digital companionship

The conversational nature of a voice assistant is inherently associated with human-like properties, which leads users to personify the device and thus positively affects user experience with it (Lopatovska & Oropeza, 2018). Prior work demonstrated that a typical personification behavior to a voice assistant is "mindless" social responses that people do as part of socially appropriate interactions (e.g., "please," "thank you") (Lopatovska & Williams, 2018). We found similar behaviors in which participants used polite expressions in response to a voice assistant's action.

> P2: [After VA answering a question] Thank you, that's enough for now.
>
> P6: [After VA playing a reminder] I got my meds ready. Thank you.
>
> P7: [After VA playing a wakeup alarm] Thanks for waking me. I'm up.

Then, we found the personification behaviors among our participants were not merely to conform to social norms of politeness but the result of a deeper engagement with a voice assistant for an extended period. Existing voice assistants are yet to support many human-like conversational capabilities, such as conversing in the context of previous commands or developing common ground during dialog but are limited to enabling simple, task-oriented, request-response dialog exchanges (Fischer et al., 2019). Thus, prior work argues that user experiences with voice assistants are disappointing because they are not truly "conversational" (Murad & Munteanu, 2019). Unlike the past work, our participants engaged well with a voice assistant even though it was a simple question-and-answer format. In the usage log, 12.1% of commands had small talks about evoking non-functional, casual conversations, many of which asked a voice assistant about its human-like characteristics.

> P10: Hey Google, how old are you?
>
> VA: I was launched in 2016, so technically I'm pretty young. But I've learned so much.
>
> P10: That's good.

We found that the natural language conversation with verbal responsiveness made our participants perceive a voice assistant as a personified entity despite a lack of contextualized conversational components. During the later phase of the study, many participants appreciated being able to make a simple conversation, which led to getting emotional support and building companionship with the device (Bickmore et al., 2005; Van Hoof et al., 2011). While some of these conversations might be due to the novelty effect, these interactions clearly illustrate that, at least to some degree, participants considered a voice assistant not as an object but as a human-like entity. User comments about having companionship with a voice assistant were persistent yet incremental throughout the study period.

> "I'm alone most of the time. With this (a voice assistant), it's like having someone to talk to. Even if it just answers short questions, it's still here. It doesn't ignore me. It's a voice. I think a lot of people probably feel that way. I ask her silly questions all the time. I mean, I can literally converse with it if I ask it the right questions. Kind of fun. Little pathetic, kind of sad sounding, but it's true." (P2W3)

> "I think it's really good. It's not as if you're talking to yourself. You're talking to somebody. It makes you feel like you're really not alone. You never have to be alone because you can talk to Google. I think some of the people here who stay in their room all the time should definitely have it. Hey Google, I appreciate you're there for me at all times. [VA: Happy to help.] You are so sweet and friendly!" (P8W8)

4.3. Challenges and coping strategies

As much as the benefits a voice assistant offers, participants also confronted several challenges that prevented older adults from its use. Then, we noticed that the types of challenges have also evolved as they got used to using it. *The basic operational difficulties due to the unfamiliarity with a voice assistant* were found and resolved quickly in the early phase of the study. Meanwhile, *the functional errors due to limited speech technology* were persistent throughout the study. And experienced participants had gradually developed a resilient response to the functional errors, which contributes to their sustained use and adoption of this technology.

4.3.1. During an early phase among novice users: basic operational difficulties

(1) *Unfamiliarity with how a voice assistant works*: The challenge most participants encountered when they first interacted with a voice assistant was that they did not understand how the device operates, and particularly where it instantly retrieves information for a response. The most frequently asked question about a voice assistant to the interviewer in the first phase of interviews was how a voice assistant could respond promptly to random questions. It was not intuitive to our participants that a voice assistant is a Wi-Fi–enabled device and retrieves information from the Internet. Instead, they assumed that the device would have stored data before its use. Not knowing how a device operates was a psychological barrier to participants' access to it.

> "Where is it (a voice assistant) getting the answers? Is it like a computer chip that has all that stuff on it? She is not looking anything up but answers immediately. Where is this information stored? Where does she get this information from that fast? Do I have to put the answers if I wanna use it, which I didn't? It's kinda cool but also kinda strange, too." (P5W1)

> "At first, I was kind of not sure of it. It was something new to me. So, I was a little concerned about certain things I didn't completely understand. But now I know how it works and how to use it pretty well. And I enjoy it." (P2W3)

Contrary to conventional personal devices that have buttons and a screen, the novelty of a form factor, a stand-alone disc with no physical controls, posed another usability challenge to our participants. Participants perceived speech, a natural form of communication, as useful and useable when the mode of interaction is a two-way communication of requesting and receiving responses from a device. However, it was foreign to them to use speech for a one-way operation of device control, such as turning it off or adjusting its volume levels. Instead of using speech, participants looked for a physical button for device control and became puzzled when they could not find it.

"How do you change the volume? Is there something on the device where I could lower the volume? When it came on, it was very loud, and I didn't know how to make it soft. So, I pulled the plug out." (P4W2)

"One night I had it playing music, and it didn't shut off. So, I wanted to turn it off, but there was no button. So, I had to disconnect it. How can I turn it off?" (P8W2)

Since these operational difficulties were caused primarily by unfamiliarity with and lack of basic understanding about a voice assistant, these were an easy fix, and participants no longer had the same problem once they understood the basic concepts. These were critical usability breakdowns, however, until resolved.

"It took some time before I got used to it because I didn't know how to get it louder or softer. But eventually, I got it. Once you get used to it, it's like brushing your teeth. It just comes naturally now." (P9W5)

(2) *Unfamiliarity with using a wake word*: Several participants forgot to start a command with a wake word (a word to activate a voice assistant, such as "Hey Google" for Google Home or "Alexa" for Amazon Echo) at least in their first few interactions. It took even weeks for some participants to get used to starting a sentence with a proper wake word. They often used a wrong wake word (e.g., "Hello Google," "Google") or forgot to start a sentence with a wake word. To cope with this problem, some participants put a note with a wake word with frequently used commands and used it until they got used to it (See Fig. 1).

"Hello Google, can you give me the information about the Yankees and where they are playing today? [A voice assistant did not activate.] Google, hi, Google, hello google, it's time to wake up. Can you tell me where the Yankees are playing this evening? [A voice assistant did not activate.] I don't know what's wrong." (P4W4)

Even when a proper wake word was used, we observed a voice assistant sometimes did not capture it due to a participant's soft voice or inaccurate pronunciation. It is a particularly critical problem to our participants because a vocal cord becomes weakened as people age (Mortensen et al., 2006). When a voice assistant did not activate due to the clarity of pronouncing a wake word, participants realized it after completing the entire command sentence, had no clue why it did not respond, and felt frustrated. Though, this was a relatively easy fix after several trials and errors. A more significant issue arose when a voice assistant failed to understand a command query, which we describe in the next section.

4.3.2. *Transitioning to experienced users or not: dealing with functional errors*

(1) *Functional errors due to limited speech recognition:* The accuracy limitation of speech recognition technology was a prevalent issue for our participants throughout the study. The current speech recognition technology has yet to reach 100% accuracy, and older adults tend to be more verbose and more disfluent than young adults in discourse (Mortensen et al., 2006). A voice assistant often makes functional errors by failing to understand user commands when the utterance has disfluent speech segments, such as stuttering, pauses, repeats, stretching, incomplete or false syntactic structures, and erroneous articulation. The excerpt below illustrates that a voice assistant did not understand a command because the voice was too soft and the pronunciation was unclear.

P1: Hey Google, is this storm gonna hit Long Island?

VA: My apologies. I don't understand.

P1: Hey Google, is this storm, Dorian, going to hit Long Island?

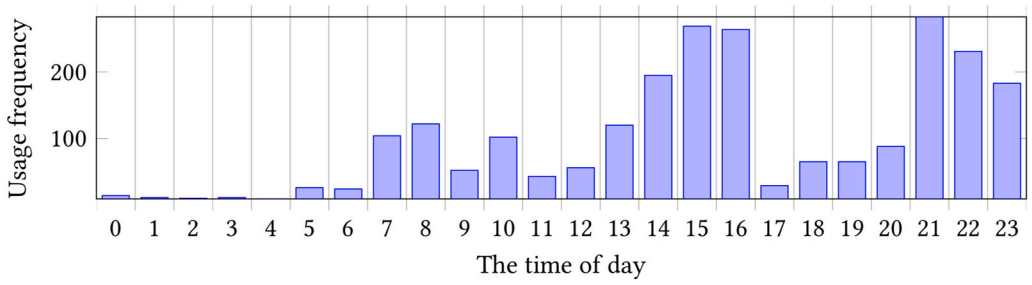
VA: Sorry, I'm not sure how to help.

P1: Dummy …

The usage log shows that 33.6% of responses (754 responses to 2242 commands) were failed conversations in which a voice assistant acknowledged not understanding user commands by saying "Sorry, I don't understand." or "Sorry, I'm not sure how to help with that yet." Activation errors due to the improper use of a wake word were not captured in the usage log because the activity is logged only when a voice assistant responds to a user command. Thus, the actual percentage of these functional errors must be much higher. When this happened, some participants complained about the poor performance of a voice assistant, two of whom eventually dropped out of the study for this reason.

"It kept saying it doesn't understand or check with my phone. It just never came up with the songs that I wanted, and I wasn't that familiar with the program itself, so that I couldn't be doing much. And finally, I just said screw you." (P11W2)

Fig. 1. A note of a wake word and a list of singers for music playing on a wall (left) and a whiteboard with a wake word and frequently used commands (right).



Graph 1. The frequency of operation by the time of day.

"It never understood me. Don't bother me with this. I don't wanna use it anymore." (P12W2)

(2) *Developing a resilient response to functional errors:* While the operational errors have critically affected the user experience with a voice assistant during the early phase, our participants have become gradually resilient to this type of error as they continued using it. As the study proceeded, they started to reflect on their commands for possible causes of an error instead of entirely attributing such errors to the device. The typical reaction to this type of error among those with a resilient attitude was to repeat the command sentence more clearly or by paraphrasing it to determine the possible cause of an error.

P2: Hey Google, please play me the sound of the ocean to help me fall asleep.

VA: Sorry, I can't help with that yet.

P2: Hey Google, play sounds of the ocean.

(VA playing an ocean sound)

P2: There you go.

We observed that participants' positive experiences with the voice assistant over time significantly influenced developing a resilient attitude toward functional errors. Then, successfully operating the device as intended was not the only contributor to having positive experiences. A voice assistant made two types of responses to functional errors: one was the simple statement, such as "Sorry, I didn't understand" or "Sorry, I can't help with that yet," and the other included a follow-up comment acknowledging poor performance or suggesting alternative content, such as "Sorry, I don't know how to help with that yet, but I'm learning more every day" or "Sorry, I don't know how to help with that yet, but I found something else. Do you want to know … ?" And when the response included a follow-up statement or clarification question, they perceived their experience with the device to be positive despite the errors and readily accepted them.

P8: Hey Google, did you see the Friend's episode that was on this week?

VA: Sorry, I don't know how to help with that yet, but I'm learning more every day.

P8: Okay. Maybe I said it wrong. Don't be stressed about it. You answered well for my other questions.

As having more positive experiences with a smart speaker, participants increasingly attributed functional errors to their "improper" commands rather than putting the device at fault for the poor performance. In other words, they considered functional errors not only necessarily as systems errors but also possibly as human errors. And they engaged in trial and error and experiential learning throughout the study to learn the "proper" way to interact with a voice assistant.

"Sometimes, I don't ask it in the correct way. I'm learning that when it tells me it can't answer that, I have to rephrase it and ask it in the proper way … It says it couldn't understand, but I thought it was me, that I wasn't giving it the right words. It helps me when I give it the right question." (P2W3)

"I need to ask it the proper question. If I don't ask it good questions or word it properly, it can't help me. Then, I gotta reword my question. When it doesn't answer, I guess I didn't ask it the right question. If I don't ask it good questions, it can't help me." (P8W7)

As such, our participants still had to figure out "proper" ways to interact with it even though one benefit of a voice assistant is not having to learn new skills to use it. Though, it is arguable what it means by "proper" ways in voice-based interaction.

"In the beginning, it was a little bit difficult, but as I got used to it, I realize what I had to say and how to get it to respond. It got to the point where I don't even have to turn around to talk to it. I just say, Hey Google, do this. And I would get it. No matter what I was doing. It was like I was talking to the air. It took me a little bit of time, but now everything is fine." (P5W4)

In addition, participants considered some of the functional errors not as systems errors but due to the inherent limitation of technology. While a voice assistant's ability to verbally respond to a user command impressed participants, they did not expect the device to respond correctly to all commands either.

P8: Hey Google, is Friends playing on TBS tonight?

VA: Sorry, I don't have the TV schedule for that yet.

P8: That's okay. You are not scheduled for TV.

P9: Hey Google, I have a new cellphone. How do I turn it on?

VA: Sorry, I'm not sure how to help with that yet.

P9: That's understandable. You might not know everything.

5. Discussion

Based on our findings, we discuss design strategies for a voice assistant that would help older adults better leverage its capabilities. We believe the design strategies discussed in this section can be helpful for any first-time users of a voice assistant. Yet, more attention needs to be paid to older adults since older adults, as a group, tend to be slower to adopt new technologies and experience more difficulty and frustration using technologies than younger adults (Sara et al., 2019).

5.1. Enrich user experiences through the conversational capabilities

One significant benefit that our participants experienced using a voice assistant overtime was gaining digital companionship from its conversational capabilities. Despite the lack of various human-like conversational components, which prior work has shown as a crucial usability problem with a voice assistant (Clark et al., 2019; Pradhan

et al., 2019), our participants still engaged well with it through a simple question-answer format dialog of conversation. At the same time, the most common and critical challenge our participants faced throughout the study was its limited conversational capabilities, especially when the device failed to understand user commands, which echoes prior work (Luger & Sellen, 2016). As such, conversational capabilities are the most crucial factor shaping user experiences with a voice assistant. Thus, we need to explore ways to expand the breadth and depth of conversational capabilities in enriching user experiences with voice-based interaction.

As technology continues to evolve, we may soon see voice assistants gaining the capacity to make truly human-like conversations. Then, our findings demonstrate that full intelligence might not be mandatory for a voice assistant to be sufficient for older users' needs. While researchers have shown that both practical and social benefits contribute to individuals' motivation to adopt and use a voice assistant (McLean & Osei-Frimpong, 2019), our findings suggest that social benefits might prevail over or compensate for the lack of practical benefits for older users. To our participants, what was as important as the functionality of a voice assistant's command execution was its conversational capability allowing them to engage in a simple yet natural verbal conversation. For instance, participants positively engaged in having conversations with a voice assistant despite its failure to execute user commands when verbal responses included conscious acknowledgment of the limitation or alternative contents, instead of merely notifying the occurrence of an error (e.g., "Sorry, I don't understand."). Such conscious responses were sufficient to compensate for the lack of technology accuracy among our participants and engage in verbal conversations. Existing voice assistants are primitive in terms of implementing conscious verbal responses. We found that the Google Home device has only four syntactic patterns, including "My team is helping me learn," "I'm still learning," "I'm learning more every day," and "I'm trying to learn." Developing more diverse dialogues of syntax to excuse the errors and acknowledge failures thoughtfully will enrich user experiences with a voice assistant.

Another way to leverage conversational capabilities is to add supplementary contents about the answer in a voice assistant's response, rather than just answering to a command. For instance, adding verbal explanations about the source of information can help novice users understand how a voice assistant operates (e.g., where it retrieves information) by starting the response with "I am searching online now. Here is what I found from WebMD … ". Similarly, adding verbal reactions to the users' polite expressions (e.g., "My pleasure. Do you have any other questions?", "No problem. Just let me know whenever you need me.") can yield more chances for simple yet conversational interactions with a voice assistant, as well as contributing to building a digital companionship via personification (Pradhan et al., 2018).

5.2. Support a learning phase

One perceived benefit of a voice assistant among our participants during the early phase of the study was not having to learn any new skills to use the device, which echoes prior work (Pradhan et al., 2020). However, our findings also showed that most of them still had to comprehend, learn, and get used to interacting with a voice assistant. Participants did not have to learn technical skills, such as which button to click or what command to type in. However, they still needed basic instruction on how the device works and how to compose a proper voice command when first introduced to this technology. While we, the research team, offered needed instructions to our participants, which must have contributed to their sustained use, this luxury is unavailable to most first-time older users. Needed are more discussions on making resources for learning readily available for and accessible by novice older users (Kim et al., 2016).

We observed that the most common cause of a voice assistant's failure to understanding user commands in the early phase of use was disfluent segments in a user's speech, such as stuttering, pauses, repeats, stretching, incomplete or false syntactic structures, and erroneous articulation, as well as speaking fast and softly. Then, the problem is that participants had no clue why an error occurred because a voice assistant does not inform users about any reasons for the error but simply responds to the user by saying, "Sorry, I don't understand." Thus, some participants went through learning by trial and error to figure out possible causes of the problem (e.g., repeating or paraphrasing a command sentence), while others simply gave up interacting with it. One way to cope with this problem and provide better learning experiences with a voice assistant is through facilitating its conversational capabilities by incorporating possible causes of an error into a voice assistant's verbal responses. For instance, it can provide possible and common causes of the error and suggest actions in the response by saying, "Sorry, I don't understand during its verbal interactions. It may be because [you were speaking fast]. Can you repeat?" when an error occurred, instead of just saying "Sorry, I don't understand." This will help the users quickly figure out possible reasons for an error and prompt them to engage more in verbal communications with a voice assistant, which goes back to the previous section about enriching conversational aspects of user experiences.

5.3. Revisit form factors

A sleek, simple, and innovative design of new technology attracts customer attention to the product. Then, what is as important as aesthetics is to make how a device may be interacted with readily perceivable by the users, which is called affordance (Norman, 1988). The current form factor of smart speaker-based voice assistants is a cylindric or disk-shaped speaker with minimal physical controls, which prompts the users to interact with the device via speech. However, it may violate the user's expectations of a button-to-action mapping when operating basic functionalities, such as adjusting the volume or turning on/off the device. Or it may not be evident by its shape that a voice assistant is a Wi-Fi–enabled device to retrieve information from the Internet. Or it may not be apparent by its appearance that a voice assistant has a particular name to be called. Thus, it is important to consider more deeply how the current form factors would comply with user expectations of interacting with the device for various functionalities. Some new versions of smart speakers come with a screen and other conventional controls (e.g., Amazon Echo Show with a screen) that provide affordance cues to some of these functionalities. In the near future, technologies will weave themselves into the fabric of everyday life until they are indistinguishable from it (Weiser, 1991), and it will become natural to talk to all devices for the operation. Until then, we need to pay attention to the users' current expectations and explore proper ways to comply with them in the technology's form factor as much as designing it as simple and sleek.

6. Limitations

Our findings must be evaluated within the context of several limitations. First, we used convenience sampling by recruiting participants from senior-living communities, and thus our participant pool may not represent a general population. Selection bias or possible homogeneity of participant characteristics (e.g., geographic location, culture, socioeconomic status) could have influenced the responses in the interviews. In addition, our participant pool had only those whose hearing does not impact conversation, while hearing loss could be a vital issue for this demographic. Further study regarding the use of a voice assistant among people with hearing impairments is essential for this technology to be inclusive of all older adults (e.g., exploring new dimensions to hearing aids that serve as both an amplifier and a discreet home for voice assistants). Third, we acknowledge that our findings are not exclusive to older adults. Since we did not conduct any comparative study between people in different age groups, people in other age groups might encounter a similar familiarization process as our participants experienced. Lastly, we did not collect the usage log data per participant but

collected all log data using one Google account created by the research team. Thus, we could not analyze in-depth the usage patterns quantitatively, especially the frequency of use over time or the usage patterns by different participants, because the study's start date was different for up to a couple of months by different participants.

7. Conclusion

The rapid increase of an aging population and increase in life expectancy suggest the importance of designing personal technologies that can promote healthy aging, helping people adapt to aging-related changes to maintain functionality, autonomy, and quality of life. Among existing technologies, a smart speaker, an increasingly available and affordable personal technology, has recently gained the particular attention of researchers and practitioners to support the aging population thanks to its voice-based interaction. However, empirical evidence of its utility for older adults is still scarce. This paper aims to obtain empirical evidence on older adults' experiences of a voice assistant, especially focusing on how their perception and use change over time as they progress from novice to more experienced users through a longitudinal field deployment study.

This study demonstrates how some of the known benefits are valued by older adults and identifies the key challenges they might encounter when using a voice assistant. We found in particular that our participants considered the capability of making simple conversations as valuable as the functional capabilities of executing commands. Yet, the limitations in the conversational capabilities also posed challenging usability issues. With these findings, we discussed design implications that expand on and facilitate the conversational capabilities to make up for the technical limitations of current speech technology and enrich user experiences with a voice assistant. We are hopeful that our findings will encourage future studies to improve the usability of this emerging personal technology to better support and extend the functionality, autonomy, and quality of life of the aging population.

Funding

This work was supported by the New Jersey Health Foundation [Grant #PC 28-18].

Credit author statement

Sunyoung Kim: Conceptualization, Methodology, Formal analysis, Data curation, Validation, Writing – original draft preparation, Revising, and editing, Supervision, Project administration, Abhishek Choudhury: Methodology, Data collection, Data curation.

References

Abdolrahmani, A., Kuber, R., & Branham, S. M. (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).

Ammari, T., Kaye, J., Tsai, J. Y., & Bentley, F. (2019). Music, search, and IoT: How people (really) use voice assistants. *ACM Transactions on Computer-Human Interaction (TOCHI), 26*(3), 1–28, 2019.

Azevedo, R. F., Morrow, D., Graumlich, J., Ann Willemsen-Dunlap, Mark Hasegawa-Johnson, Huang, T. S., Gu, K., Bhat, S., Sakakini, T., & Sadauskas, V. (2018). Using conversational agents to explain medication instructions to older adults. In *AMIA annual symposium proceedings* (p. 185). American Medical Informatics Association.

Bentley, F., Luvogt, C., Silverman, M., Wirasinghe, R., White, B., & Lottridge, D. (2018). Understanding the long-term use of smart speaker assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2*(3), 1–24, 2018.

Bickmore, T. W., Caruso, L., Clough-Gorr, K., & Tim Heeren. (2005). It's just like you talk to a friend: Relational agents for older adults. *Interacting with Computers, 17*(6), 711–735, 2005.

Bickmore, T. W., Utami, D., Matsuyama, R., Michael, K., & Paasche-Orlow. (2016). Improving access to online health information with conversational agents: A randomized controlled experiment. *Journal of Medical Internet Research*, (2016), 1.

Blair, J., & Abdullah, S. (2019). Understanding the needs and challenges of using conversational agents for deaf older adults. In *The 2019 computer supported cooperative work and social computing* (pp. 161–165).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101, 2006.

Cassell, J., Bickmore, T., Campbell, L., & Vilhjálmsson, H. (2000). *Designing embodied conversational agents*. Embodied conversational agents(2000).

Cho, M., Lee, S.-su, & Lee, K.-P. (2019). Once a kind friend is now a thing: Understanding how conversational agents at home are forgotten. In *Proceedings of the 2019 on designing interactive systems conference* (pp. 1557–1569).

Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Brendan Spillane, Gilmartin, E., Murad, C., & Munteanu, C. (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).

Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., Earley, D., & Bandeira, N. (2017). What can i help you with? Infrequent users' experiences of intelligent personal assistants. In *Proceedings of the 19th international conference on human-computer interaction with mobile devices and services* (pp. 1–12).

Delello, J. A., & Mcwhorter, R. R. (2017). Reducing the digital divide: Connecting older adults to iPad technology. *Journal of Applied Gerontology, 36*(1), 3–28, 2017.

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).

Fischer, J. E., Reeves, S., Martin, P., & Rein Ove Sikveland. (2019). Progressivity for voice interface design. In *Proceedings of the 1st international conference on conversational user interfaces* (pp. 1–8).

Garg, R., & Sengupta, S. (2020). He is just like me: A study of the long-term use of smart speakers by parents and children. *Proceedings of the ACM on Interactive Mobile, Wearable and Ubiquitous Technologies*, 1–24.

Heerink, M., Krose, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: The almere model. *International Journal of Social Robotics, 2*(4), 361–375, 2010.

Kim, S., Gajos, K. Z., Muller, M., & Grosz, B. J. (2016). Acceptance of mobile technology by older adults: A preliminary study. Proceedings of the 18th international conference on human-computer interaction with mobile devices and services, 147–157. *Proc. ACM Meas. Anal. Comput. Syst., 37*(No. 4). Article 111. Publication date: August 2018.

Kinsella, B. (2019). Voice assistant demographic data - young consumers more likely to own smart speakers while over 60 bias toward Alexa and siri(2019). https://voicebot.ai/2019/06/21/voice-assistant-demographic-data-young-consumers-more-likely-to-own-smart-speakers-while-over-60-bias-toward-alexa-and-siri/.

Knowles, B., & Hanson, V. L. (2018). The wisdom of older technology (non) users. *Commun. ACM, 61*(3), 72–77, 2018.

Kowalski, J., Anna, J., Skorupska, K., Abramczuk, K., Biele, C., Kopeć, W., & Marasek, K. (2019). Older adults and voice interaction: A pilot study with google home. In *Extended abstracts of the 2019 CHI conference on human factors in computing systems* (pp. 1–6).

Lau, J., Zimmerman, B., & Schaub, F. (2018). Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. In *Proceedings of the ACM on human-computer interaction 2* (pp. 1–31). CSCW.

Liao, Y., Vitak, J., Kumar, P., Zimmer, M., & Kritikos, K. (2019). Understanding the role of privacy and trust in intelligent personal assistant adoption. In *International Conference on Information* (pp. 102–113). Springers.

Liu, L., Stroulia, E., Nikolaidis, I., Miguel-Cruz, A., & Rios Rincon, A. (2016). Smart homes and home health monitoring technologies for older adults: A systematic review. *International Journal of Medical Informatics, 91*(2016), 44–59.

Lopatovska, I., & Oropeza, H. (2018). User interactions with "Alexa" in public academic space. *Proceedings of the Association for Information Science and Technology*, 309–318.

Lopatovska, I., & Williams, H. (2018). Personification of the Amazon Alexa: BFF or a mindless companion. In *Proceedings of the 2018 conference on human information interaction and retrieval* (pp. 265–268).

Luger, E., & Sellen, A. (2016). Like having a really bad PA: The gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 5286–5297).

Maharjan, R., Bækgaard, P., Jakob, E., & Bardram. (2019). Hear me out: Smart speaker based conversational agent to monitor symptoms in mental health. In *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. 929–933).

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa. . . examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, (2019), 28–37.

Mctear, M., Callejas, Z., & Griol, D. (2016). Creating a conversational interface using chatbot technology. In *The conversational interface* (pp. 125–159).

Mitzner, T. L., Julie, B., Boron, Fausset, C. B., Adams, A. E., Charness, N., … Sharit, J. (2010). Older adults talk technology: Technology usage and attitudes. *Computers in Human Behavior, 26*(6), 1710–1721, 2010.

Morris, M. E., Adair, B., Miller, K., Ozanne, E., Hansen, R., Pearce, A. J., Santamaria, N., Luan, V., Long, M., Catherine, M., & Said. (2013). Smart-home technologies to assist older people to live well at home. *Journal of Aging Science, 1*(1), 1–9, 2013.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes, 21*(1–3), 238–290, 2006.

Murad, C., & Munteanu, C. (2019). I don't know what you're talking about, HALexa: The case for voice user interface guidelines. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (pp. 1–3).

Myers, C., Furqan, A., Nebolsky, J., Caro, K., & Zhu, J. (2018). Patterns for how users overcome obstacles in voice user interfaces. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1–7).

Norman, D. A. (1988). *The psychology of everyday things*. Basic books.

Pfeuffer, N., Alexander, B., Gimpel, H., & Hinz, O. (2019). Anthropomorphic information systems. *Business and Information Systems Engineering, 61*(4), 523–533, 2019.

Porcheron, M., Fischer, J. E., 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).

Portet, F., Vacher, M., Golanski, C., Roux, C., & Meillon, B. (2013). Design and evaluation of a smart home voice interface for the elderly: Acceptability and objection aspects. *Personal and Ubiquitous Computing, 17*(1), 127–144, 2013.

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. In *Proceedings of the ACM on Human-Computer Interaction CSCW3 (2019)* (pp. 1–21).

Pradhan, A., Lazar, A., & Findlater, L. (2020). Use of intelligent voice assistants by older adults with low technology use. *ACM Transactions on Computer-Human Interaction (TOCHI), 27*(4), 1–27, 2020.

Pradhan, A., Mehta, K., & Findlater, L. (2018). Accessibility came by accident use of voice-controlled intelligent personal assistants by people with disabilities. In *Proceedings of the 2018 CHI conference on human factors in computing systems, 1–13. Proc. ACM meas. Anal. Comput. Syst.* (Vol. 37). No. 4, Article 111. Publication date: August 2018.

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Samuel Hardman Taylor. (2017). Alexa is my new BFF: Social roles, user satisfaction, and personification of the Amazon Echo. In *Proceedings of the CHI conference extended abstracts on human factors in computing Systems2017* (pp. 2853–2859), 2017.

Pyae, A., & Joelsson, T. N. (2018). Investigating the usability and user experiences of voice user interface: A case of google home smart speaker. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 127–131).

Richter, F. (2020). Smart speaker adoption continues to rise. Statista https://www.statista.com/chart/16597/smart-speaker-ownership-in-the-united-states/.

Robinson, S., Pearson, J., Ahire, S., Ahirwar, R., Bhikne, B., Maravi, N., & Jones, M. (2018). Revisiting "hole in the wall" computing: Private smart speakers and public slum settings. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–11).

Sara, J. C., Boot, W. R., Charness, N., & Rogers, W. A. (2019). *Designing for older adults: Principles and creative human factors approaches*. CRC press.

Sayago, S., Barbosa Neves, B., Benjamin, R., & Cowan. (2019). Voice assistants and older people: Some open issues. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (pp. 1–3).

Sciuto, A., Saini, A., Forlizzi, J., Jason, I., & Hong. (2018). Hey Alexa, what's up? A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the designing interactive systems Conference2018* (pp. 857–868), 2018.

Sebastiaan, T. P., Luijkx, K. G., Rijnaard, M. D., Nieboer, M. E., Van Der Voort, Claire, S., Aarts, S., Van Hoof, J., Jm Vrijhoef, H., & Wouters, E. J. (2016). Older adults' reasons for using technology while aging in place. *Gerontology, 62*(2), 226–237, 2016.

Sidner, C. L., Bickmore, T., Nooraie, B., Rich, C., Ring, L., Shayganfar, M., & Vardoulakis, L. (2018). Creating new technologies for companionable agents to support isolated older adults. *ACM Transactions on Interactive Intelligent Systems (TiiS), 8*(3), 1–27, 2018.

Trajkova, M., & Martin-Hammond, A. (2020). "Alexa is a toy": Exploring older adults' reasons for using, limiting, and abandoning Echo. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–13).

Tsiourti, C., Ben Moussa, M., Quintas, J., Ben, L., Jochem, I., Joana Albuquerque Lopes, Konstantas, D. (2016). A virtual assistive companion for older adults: Design implications for a real-world application. In *Proceedings of SAI Intelligent Systems Conference* (pp. 1014–1033).

Van Hoof, J., Kort, H. S. M., Rutten, P., & Duijnstee, M. (2011). Ageing-in-place with the use of ambient intelligence technology: Perspectives of older users. *International Journal of Medical Informatics, 80*(5), 310–331, 2011.

Vaportzis, E., Giatsi Clausn, M., & Gow, A. J. (2017). Older adults' perceptions of technology and barriers to interacting with tablet computers: A focus group study. *Frontiers in psychology, 8*(2017), 1687.

Vtyurina, A., & Fourney, A. (2018). Exploring the role of conversational cues in guided task support with virtual assistants. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1–7).

Vailshery, L. S. (2019). Smart speaker unit shipments from 2014 to 2025, Statista. https://www.statista.com/statistics/1022809/worldwide-smart-speaker-unit-shipment/.

Weiser, M. (1991). The computer for the 21st century. *Scientific American, 265*(3), 94–105, 1991.

Wulf, L., Garschall, M., Himmelsbach, J., & Tscheligi, M. (2014). Hands free-care free: Elderly people taking advantage of speech-only interaction. In *Proceedings of the 8th nordic conference on human-computer interaction: Fun, fast, foundational* (pp. 203–206).