Designing Voice-First Ambient Interfaces to Support Aging in Place

Andrea Cuadra
Stanford University
Stanford, CA, USA
apcuad@stanford.edu

Jessica Bethune
Independent Researcher
Seattle, WA, USA
jessica.bethune@gmail.com

Rony Krell
Cornell Tech
New York, NY, USA
rk453@cornell.edu

Alexa Lempel
Cornell Tech
New York, NY, USA
al2444@cornell.edu

Katrin Hänsel
Yale University
New Haven, CT, USA
katrin.hansel@yale.edu

Armin Shahrokni
Hackensack Meridian Health
Neptune, NJ, USA
armin.shahrokni@hmhn.org

Deborah Estrin
Cornell Tech
New York, NY, USA
destrin@cornell.edu

Nicola Dell
Cornell Tech
New York, NY, USA
nixdell@cornell.edu

ABSTRACT

We focus on the stories of five older adults who became voice assistant users through our study, and with whom we speculated about future interfaces through two design probes, one for health data reporting and one for positive reminiscing. We delivered a voice-first ambient interface (VFAI) to each participant, and closely observed participants' journeys through periodic themed interviews (16 hours, 21 minutes of transcribed recordings), usage log reviews (4,657 entries), and phone and text support. Participants' lived experiences impacted their perceptions and interactions with their VFAI, fueling rich insights about how to design for diverse needs. For example, while one participant saw increased potential in the VFAI after interacting with the design probe for health data reporting, another was skeptical of using it to communicate with her doctor. We contribute an in-depth exploration of VFAIs to support aging in place, implications for design, and areas for future work for tailoring VFAIs towards enabling continuity of care in people's homes.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in ubiquitous and mobile computing*; *Empirical studies in HCI*; *Field studies*; User studies; **Participatory design**; User centered design; Contextual design; **Sound-based input / output**; **Natural language interfaces**.

DIS '23, July 10–14, 2023, Pittsburgh, Pennsylvania, USA

ACM ISBN 978-1-4503-9893-0/23/07…$15.00
https://doi.org/10.1145/3563657.3596104

KEYWORDS

Older adults; smart speakers; voice assistants; Alexa; home health; wellbeing; internet of things; empirical study; field study; interviews; prototyping/implementation; qualitative methods; design probes; voice-first ambient interfaces; inclusive design; aging in place

ACM Reference Format:
Andrea Cuadra, Jessica Bethune, Rony Krell, Alexa Lempel, Katrin Hänsel, Armin Shahrokni, Deborah Estrin, and Nicola Dell. 2023. Designing Voice-First Ambient Interfaces to Support Aging in Place. In *DIS '23: ACM SIGCHI Conference on Designing Interactive Systems, July 10–14, 2023, Pittsburgh, Pennsylvania, USA.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3563657.3596104

1 INTRODUCTION

The proportion of older adults in the US population is larger than ever before and the growth trend is expected to continue [18], widening the need to support aging-in-place [45]. Many older adults desire to live independently at home for as long as possible, suggesting potential opportunities for technological innovation [35, 67]. As we age, our bodies begin to function differently: we may not remember things like we used to, have shaky hands, and/or need glasses to read text. We may lose our partners or friends, and could end up living alone, creating a sense of isolation that can be detrimental to health and wellbeing [24]. We may become dependent on others, such as family caregivers or home care workers. In the face of these challenges, new technologies, such as VFAIs, could help empower older adults wishing to age in place, enabling them to maintain independence longer while supporting their health and wellbeing [6, 56, 90].

Voice assistants (e.g., Amazon's Alexa, Google's Assistant, or Apple's Siri) are VFAIs that exist in devices that connect to the Internet, such as smart speakers, watches, and mobile phones. Voice assistants are meant to be easy to use because people simply talk to them to get a response. It is estimated that, by 2024, more than eight billion voice assistant devices will be in use globally [5]. Industry players have already begun investing in voice app startups for

health [1, 4] and wellbeing [2]. For older adults who may be digitally low-literate, unable to travel to seek care, have motor or visual impairments [15], or who are otherwise unable to access existing forms of care, voice assistants could be a portal for accessing health and wellbeing services at home. Moreover, voice assistants have the potential to help healthcare providers and wellbeing services effectively reach older patients in their homes, at scale.

However, VFAIs designed for and with older adults are scarce and understudied. This may be due to older adults being frequently excluded from research and design of mainstream technologies, such as voice assistants [11, 72, 85], resulting in technological products that do not address their needs [30, 31]. Despite voice assistants' promise for supporting home health and wellbeing, investigation of the needs of older adults has been insufficient [78]. Recently, more studies have investigated how VFAIs are used in older adults' homes [51, 68, 92] or how they may be designed for older adults' health information seeking needs [59]. However, to the best of our knowledge, no studies have combined extended, in-situ use of VFAIs with the deployment of design probes to explore how VFAIs may support aging in place. VFAIs designed with and for older adults may help to bridge the digital divide in the provision of healthcare and wellbeing services.

This paper presents an in-depth, research-through-design [91] exploratory field study of our participants' (n=5) journeys becoming Echo Show users and their interactions with two design probes to support aging in place, one inspired by the promise of VFAIs for home health and the other for wellbeing. We chose research-through-design as a method, because it uncovers unknown unknowns, leading to human-centered innovations that respond to real needs. To conduct the study, we delivered Amazon Echo Shows in person to older adults who lived alone, and introduced participants to their new VFAI. We closely observed their journey for approximately two months by conducting periodic themed interviews (16 hours, 21 minutes of transcribed recordings), reviewing their usage logs (4,657 entries), and providing phone and text support. We deployed design probes (prototype voice-first multi-modal apps) to participants' Echo Shows before the second and third interviews. One design probe was a voice-based geriatric assessment developed in partnership with a hospital. The other design probe engaged participants in positive reminiscing by using questions from StoryCorps[1] [3] as prompts. **The goal of this study was to explore designing VFAIs to support aging in place by using design probes to understand VFAIs' potential to support health and wellbeing for older adults living independently.**

We found that the design probes helped participants see the possibilities of using Alexa for health and wellbeing in a way that they had not imagined or mentioned before using them, and formulate detailed opinions surrounding these speculated futures. We demonstrate a range of conflicting perspectives about these futures—while some saw them as valuable, others expressed doubts. Our probes helped us identify important design considerations, such as the need to address participants' perception that information collected via a voice-based health questionnaire would also need to be consumed as audio, which could "bother" their doctors. We also highlight several examples of how participants' lived experiences impacted their interactions with Alexa. For example, one participant who occasionally felt judged by people valued Alexa as a non-judgmental companion, and a participant who was a teacher valued Alexa as an interactive agent she could teach things to. In addition, engaging with the design probes empowered participants to contribute ideas for improvements and for other interfaces that would be useful to them. Our in-depth observations of each participant's unique experience fuel rich insights for the design of VFAIs to support older adults' health and wellbeing.

As a whole, we contribute an in-depth exploration of VFAIs to support aging in place, implications for design, and areas for future work for tailoring VFAIs to older adults. These findings are needed to inform future research for technologies to support continuity of care in older adults' homes.

2 RELATED WORK

As we now describe, prior research has examined how VFAIs are used by older adults. Efforts have been made to investigate VFAIs' role as a technological solution that may support aging in place, in particular via home health, and on their role as companions that may help improve wellbeing. Our paper extends these findings by characterizing five diverse users' experiences becoming familiarized with Echo Shows, and prompting them to imagine alternatives with us in a deeply human-centered manner. We ground our speculative conversations on empirical observations from participants' interactions with the devices.

2.1 VFAI use by older adults

Despite the unique promise VFAIs present for older adults, VFAI research focused on older adults has only recently started to proliferate. In 2019 at the first ACM Conversational User Interfaces conference, Sayago et al. [72] encouraged scholars to look into many open issues related emerging from how little attention had been paid to this important research area. For example, in the same year, a systematic literature review of research published in the Association of Computing Machinery Digital Library only found 16 articles addressing the perception and use of VFAIs by older adults as opposed to 1581 records when age-related keywords were removed [78]. This is a missed opportunity, as an analysis of VFAI usage logs by Oh et al. [64] found that older adults used the VFAI more when compared to younger adults.

In 2020, Pradhan et al. [68] conducted a general-use study with similar methods to ours, a 3-week field deployment of the Amazon Echo Dot in the homes of seven older adults, and found consistent usage for finding online information, much of this information was health-related. In another study, also with similar methods to ours, Kim and Choudhury [51] found that over time older adults felt less worried about making mistakes and enjoyed the digital companionship as they got used to using VFAIs. Similarly, Zubatiy et al. [92] found that VFAIs empowered older adults with mild cognitive impairments and their care partners. We draw from the deployment methodology employed in these studies, but dive deeper by utilizing design probes to speculate about and consider potential futures together.

[1]StoryCorps is a non-profit organization centered around storytelling and popularized by National Public Radio. More information is available at https://www.npr.org/

In their study, O'Brien et al. [62] identified five major themes for older adults' use of VFAIs: (1) entertainment; (2) companionship; (3) home control; (4) reminders; and (5) emergency communication. Out of these, though not explicitly stated in the paper, the first two may be related to wellbeing due to their ability of addressing emotional and social needs [81], and the last two to home health—such as by providing verbal reminders for medications, or supporting verbal requests for help in the case of a medical emergency (e.g., a fall). Ennis et al. [35] developed a VFAI-based innovation to support ambient assistive living and received promising feedback. However, despite their promise, many older adults abandon the use of these devices, citing difficulty in finding valuable uses, beliefs about the lack of essential benefits provided, or challenges with use in shared spaces [80]. There are several studies that seek to explain this. Prior work exploring technology use by older adults has suggested the need for "additional support and training and/or alternative interface designs" [36]. A more recent scoping review by Arnold et al. [7] found that most of the existing literature focuses on technology usability, acceptance, and common uses. Moreover, none of the included studies evaluated the effectiveness of the technologies' ability to improve the management of health conditions or to facilitate the functional capacity of older adults. Despite the emergence of new literature in the space [16, 25, 32, 38, 76, 77, 81], how to design VFAIs in a truly inclusive and acceptable way for those who may stand to benefit the most from them is still an open question [75]. These trends and findings informed our interview guides, participant support structure, and research direction, allowing us to create an adequate space for design speculation to address the scarcity of information about how to design VFAIs that are truly inclusive of older adults' needs and preferences.

2.2 VFAIs for health and wellbeing

Healthcare literature increasingly recognizes the need for technological solutions to support older adults' healthcare needs [6, 56, 67, 74, 90], an essential component for supporting aging in place. Abdi et al. [6] specifically cite VFAIs as one of eight emerging technologies that could potentially be used to meet older people's needs in various care and support domains. Several researchers have investigated the use and perception of conversational interfaces for health information seeking [14, 42, 43, 47, 59, 61]. A highly relevant study was conducted by Harrington et al. [43] with Black older adults from lower socioeconomic environments who may particularly benefit from using VFAIs for health information seeking, and found that many older adults in their sample struggled with the mechanics of initiating interaction and wording questions in a way that would support relevant responses. Their participants drew parallels between navigating speech with conversational technologies and cultural code-switching done in their everyday lives, highlighting the urgency to design these interfaces inclusively. In an adjacent line of research, Chen et al. [20] explored VFAI care delivery and quality of life enhancements for older adults as a collaborative task between patients and providers, highlighting more support for health data reporting (what our first design probe does) as a major application for VFAIs. Furthermore, we are also concerned with wellbeing, which is related to health—defined by the Oxford Languages dictionary as *the state of being comfortable, healthy, or happy*. Thus, we build a second design probe that combines VFAIs strength as companions with questions from StoryCorps that elicit positive reminiscing through storytelling, and describe related work on companionship and positive reminiscing below.

2.2.1 Companionship. Although computer use alone has not been associated with wellbeing [33, 52, 87], prior research has found that some older adults enjoy VFAIs' companionship [21, 23, 51, 62, 66]. Healthy cognition is an important aspect of wellbeing, and factors such social disengagement have been directly linked to cognitive decline [8]. Because of this, social relationships with VFAIs could be leveraged to support older adults' wellbeing.

The role of VFAIs as companions has been noted by many. In an analysis of Amazon reviews about VFAI use by older adults, O'Brien et al. found that one of the categories older adults use VFAIs for is companionship [62]. More deeply examining the ontological categorization of VFAIs, Pradhan et al. [66] found that participants desiring companionship were more likely to value the VFAI as a companion, and that the categorization of the device between companion and object was fluid. In the same vein of research, Chung et al. found that older adults tended to personify the agent more than younger counterparts by using polite words such as "grateful", viewing it more as a companion [21]. Moreover, Kim and Choudhury [51] found participants built digital companionship as they became more experienced with VFAIs. Corbett et al. [23] made a call for more research in this area, as their mini review of literature covering older adults use of VFAIs and its relationship to social isolation and loneliness suggested a promising role for VFAIs as companions to reduce loneliness. These findings are also supported by research on similar technological agents [54, 70, 83].

2.2.2 Positive reminiscing. We use a positive reminiscing design probe to speculate about how VFAIs can support wellbeing, because of the existing evidence about the potential of positive reminiscing to improve wellbeing. In a systematic review, Hsieh and Wang [44] found that reminiscence therapy has resulted in statistical significantly decrease in depression. O'Rourke et al. [65] found that prosocial reminiscence functions, such conversing or teaching others, appear to have an indirect association with well-being. Finally, Cappeliez et al. [19] found that the vast majority of narrative reminiscences occurred in the context of positive emotions, either eliciting, maintaining, or even amplifying positive feelings.

2.3 Designing with older adults

We rely on functional prototypes deployed to participants' devices as our design probes to be able to mimic real-life, naturalistic interactions. This decision is informed by prior work by Lazar et al. [53] observing the formation of a makerspace in a retirement community. They found that "to answer questions about their preferences for a makerspace, participants had to first form a sense of what a makerspace is or is not for." Without this information, ideas were too broad or too narrow. Lindsay et al. [55] similarly identified how older adults might have challenges in brainstorming invisible technological futures. In our study, without knowing the specific capabilities and limitations of Echo Shows, participants would have

difficulty formulating detailed preferences or coming up with well-scoped ideas for VFAIs. Moreover, given that ease-of-use and usefulness are significant variables in predicting technology acceptance [60], it is important that these interactions are studied in a realistic setting where usability issues can be noted. To investigate specific interface designs, some have done participatory design workshops with scenarios [59], and others have created and tested prototypes that rely on scripted utterances [61]. To advance early-stage investigations that do not fully cover the range of limitations that come into play in practice, we build functional prototypes as our design probes for this exploration.

2.4 Combating stereotypes surrounding older adult technology use

In response to harmful stereotypes made about older adults, Rogers and Marsden [71] called for us to move beyond the rhetoric of compassion to one of empowerment. Furthermore, Vines et al. [84] found that discourse is often framed in problematic ways that represent common stereotypes around old age, and call for more-individualized and more-contextualized approaches. Similarly, Harrington et al. [41] recommend embracing the full humanity of research participants. Older adults who adopt information communication technology have been found to value their technology activities as important in their lives [86]. Responding to this research, we focus this work on the user journeys of only a few older adults (to be able to provide more-individualized and more-contextualized attention) from demographics that have been historically underrepresented in the use and design of technology (see Section 3.1), and supply the smart speaker devices and training on their use as part of the research.

3 METHOD

We conducted an IRB-approved (protocol #1912009271) in-the-home study with five older adults living alone, recruited via local senior centers in New York City. The interviews were all conducted by the first author remotely via phone or Zoom while participants were in their homes. The design probes were remotely deployed to participants' devices. We selected Alexa smart speakers with screens (Amazon Echo Shows) as the voice assistant devices for the study, because of the promise of multi-modal interfaces to bridge interaction gaps that could enhance inclusion. The interviewer's positionality as a scholar with intersecting marginalized identities likely helped deepen the conversations with participants [63], and establish an ethos of care.

3.1 Participants

We recruited five older adults (four women, one man) between the ages of 62 and 85, with varying degrees of technical familiarity (see Table 1 for more details). To recruit participants, we sought people who belonged to older adult communities [12, 69], senior centers. Senior centers are community centers designed to make older adults feel supported, and happy—they bring older adults together for a variety of services and activities designed to enhance their quality of life [9]. We gave a short presentation about the study during a Zoom meeting with many senior center directors, and sent each director a flyer with details about our study to share with their members. Some directors got back to us with prospective participants' names and phone numbers. We then called each prospective participant, explained the details of the study, and answered any questions they had. They were expecting our call, because their senior center director had told them about us. If a person was interested in participating, we arranged a time to drop off the device, and obtained consent. Note, higher prevalence of technology use in older adults has been associated with five characteristics: younger age, male sex, white race, higher education level, and being married (all p values <.001) [36]. We aimed to focus on older adults who have been historically underrepresented [41], so we recruited participants who did not have more than two of the five characteristics associated with higher prevalence of technology use (i.e., none of our participants were married, the younger ones were not white, and the only man was low-literate). All participants lived by themselves independently and had WiFi in their homes. This study responds to Dix [34] who argues for the value of small-scale studies "as we move from a small number of applications used by many people to a 'long tail' where large numbers of applications are used by small numbers of people," and Vines et al. [84] who suggest critical engagement with an individual's context as a strategy to combat common stereotypes that prevail in the literature. Keeping the number of participants at five allowed us to create the deep, personalized engagements we sought. We characterize the individuality of our five participants, highlighting the diversity of older adults' uses of, and reactions to, new technologies, and provide rich descriptions of each participant, including more demographic details, in Section 4.1.

3.2 Procedure and Materials

We dropped off multi-modal smart speakers at each participant's home and allowed them to get familiarized with the device for at least two weeks before the first remote interview. This gave participants enough time to explore the device in a way that mimics a real-life situation (e.g., where they may have received the device as a gift). The devices were Amazon Echo Shows (second generation), with a 10.1 inch high-definition smart display with Alexa. During drop-off, we helped participants setup the devices with their home WiFi and experientially showed them examples of common uses (e.g., weather forecasts, music, and information retrieval), explained how to mute the device if they did not want it to be "listening," and answered any questions they had. After they had spent at least two weeks with the installed device, we conducted three audio-recorded, hour-long, semi-structured interviews with each participant (see Table 2 for a summary of each interview's content). The interviews were spaced at least five days apart from one another, and recorded for transcription (see Table 1 for the exact number of days between interviews). The interviews also served as ways to answer participants' questions. Interviews were conducted either via phone calls or Zoom video conferencing meetings (depending on the participants' preferences and abilities), and were themed around the following topics: (I1) familiarization, (I2) health, and (I3) wellbeing. For the latter two interviews, we employed a research through design approach [91] by creating two design probes—one for health data reporting, and the other for positive reminiscing–as design

Table 1: Summary of participant demographics, home computer (PC) ownership, interview timelines, and usage trends.

Name	Gender (& age)	Race	PC	Zoom	Device location	I1 (days from 1st use)	I2 (days from I1)	I3 (days from I2)	Total weeks	Usage trend	Mean daily interactions
Travis	M (67)	Black	No	No	Bedroom	40	20	28	12.6	↑	7.64
Gilda	W (82)	White	Yes	Yes	Home office	15	6	20	5.9	↑	3.54
Betty	W (85)	White	Yes	Yes	Living room	22	8	25	8.0	↓	4.39
Theresa	W (85)	White	Yes	Yes	Home office	21	17	19	8.1	↑	5.91
Maria	W (62)	Other	No	Yes	Living room	20	14	26	8.6	↓	17.40

To preserve anonymity, names shown are not participants' real names. We determined trends by comparing the number of interactions in the first half of the total number of days with the device to those in the second half; ↑ indicates increased usage and ↓ indicates decreased usage. Interaction counts exclude "Alexa" or "echo" single word transcripts, or recorded audio snippets labeled by Amazon as "Audio was not intended for Alexa" (or similar).

probes to engage participants in in-depth discussions and ideation about the potential of using Alexa for health and wellbeing.

3.2.1 Design probe implementation. The design probes were deployed to, and ran from, participants' devices using dedicated email and Amazon accounts for each device. We built the design probes using Amazon's Alexa Skills Kit (ASK), which is a compilation of open-sourced Alexa application programming interfaces and tools to develop voice apps. We deployed the design probes to participants' devices remotely by inviting the devices' dedicated email accounts to be Beta testers, and accepting those invitations via the devices' Amazon accounts (which used the same email address). We shared the specific phrase to run the design probes during their respective interviews. We explicitly asked them to use the design probe for health data reporting once outside of interview times. We tracked usage of the voice assistant and design probes by reviewing the interaction history on the devices' accounts.

3.2.2 Familiarization interview. The first interview was focused on establishing grounding with participants as they became familiarized with the Echo Shows. In this interview, we sought to understand the context in which participants had been using the device, and what they perceived the voice assistant's strengths and challenges to be. Knowing the strengths they perceived allowed us to tailor recommendations for more potential uses, and knowing the challenges allowed us to provide additional support. In addition, we relied on usage logs to dive deeper into participants' experiences with the device by asking about interactions that they might not have brought up during the interview. In summary, the familiarization interview served many purposes: 1) to provide additional training, as recommended in existing literature [23, 29, 87]; 2) to establish rapport between the interviewer and each participant; and 3) to document the participants' perceptions and experiences with the smart speaker before introducing our design probes. We started each of the following interviews with a general update, asking: How are you? How has it been going with the Alexa device? Did you try anything new from last time? Has anything changed? Has anything surprised you?

3.2.3 Interview with the design probe for health data reporting. The second interview was centered around uses for home health. We asked questions such as: What do you do to stay healthy at home? What do you wish you did better to stay healthy at home? What sorts of concerns would using Alexa for home health introduce? How would you feel about using the Alexa device to communicate with your doctor? We chose health data reporting as an appropriate application for voice assistant technology, because this was a need raised by healthcare practitioners across multiple institutions in our exploratory phase (e.g., [28]). The design probe for reporting health data to doctors and nurses (see Figure 1) was a section of an established geriatric assessment [74] adapted from written to spoken format. We chose these questions, because they are used by our collaborators' geriatric services department to measure patients' fitness for receiving cancer treatment, a realistic and relevant future application for a VFAI. For this interview, we introduced an imaginary scenario and asked participants to interact with the design probe. At the end of the interview, we asked participants to use the design probe once on their own before the next interview. The design probe we developed was called *My Care Questionnaire*, and asked how much users had been limited in seven activities of daily living (ADLs) by their health condition, as follows:

Participant: Alexa, open My Care Questionnaire.
Alexa: Welcome to My Care Questionnaire. Your answers will help your medical team provide more comprehensive care for you, which may improve your outcome. Are you ready? You can say yes or no.
Participant: Yes.
Alexa: All right. Let's get started. How much have the following activities been limited by your health condition? For each activity you can say, not limited, limited a little, or limited a lot. The first activity is bathing.
Participant: Not limited.
Alexa: How about dressing? …

For each ADL, a screen was also displayed (Figure 1) with the answer options and an image for the activity: bathing, dressing, grooming, feeding, walking inside the home, walking outside the home, or bladder and bowel control. At the end of the interaction, the VFAI said, "Your questionnaire has been successfully completed! I will send your responses to your care team. If you have any questions, reach out to your doctor's office. Have a good day!"

3.2.4 Interview with the design probe for positive reminiscing. The third interview explored the use of the voice assistant for wellbeing. Before introducing the design probe, we asked questions such as: What do you do to stay at peace? What makes you feel proud? What's your attitude towards loneliness? Do you like sharing stories about your life? How about listening to stories? With whom?

Table 2: Summary of interviews.

Interview 1 (I1): Familiarization	Interview 2 (I2): Home Health	Interview 3 (I3): Wellbeing
Alexa strengths and challenges	General update	General update
Questions from usage logs	Design probe #1	Design probe #2
Brainstorm potential uses	Homework: try design probe once on their own	Reflection of the full experience

Figure 1: The design probe for health data reporting. Left: Amazon Echo Show 10 displaying the bathing ADL. Right: thumbnail images for the remaining 6 ADL screens.

Why? Then, we dove deeper via a prototype we developed to facilitate human-to-human or human-to-machine positive reminiscing by providing story prompts from StoryCorps [3], a non-profit organization whose mission is to record, preserve, and share the stories of people in the U.S. from all backgrounds and beliefs. Positive reminiscing can boost happiness [17], improving wellbeing. We chose to employ a design probe that was not strictly within the health domain as a way to explore a use case that deviates from discourses surrounding health economics or deficit [84]. The design probe can be used in group settings or by oneself. The invocation phrase for the prototype was "load a StoryCorps question." This phrase would start the app, which would set the stage by telling users "great questions are on the way," and encouraging them to find a comfortable position, and a recording device if they wanted to record the conversation.[2] Since we were recording the interview, there was no need to record on another device. Once a participant told the app that they were ready, a question would appear on the screen and the voice assistant would also say the question out loud. We used three questions: 1) "What's one of your happiest memories?" 2) "What was your childhood like?" and 3) "What are some of the most important lessons you've learned in life?"[3] Each question would appear individually and the next question would be shown when the participant requested it. In the software backend, we made Alexa play silent audio so that it stayed active (and did not quit the voice app) while users shared their stories. When the three questions have been asked, if the participant requested another question the app would say, "There are no more questions left for today. Come back tomorrow for more."

3.3 Data Analysis

We closely reviewed usage log transcripts (before interviews and at the end of the study), and analyzed interview transcripts. We also considered notes from interactions during device drop off in participants' homes, the initial orientation, and installation process when relevant (as in Gilda's description in Section 4.1).

[2]The design probe required users to use their own device because at the time of the study, Alexa did not allow third-party developers a straighforward way to access voice recordings. As a result, to track usage of our design probe, we had to rely on the automated transcripts generated by Alexa.

[3]These questions were obtained from StoryCorps's list of Great Questions: https://storycorps.org/participate/great-questions/

3.3.1 Usage log transcripts. We recorded 4,657 usage log entries from the research accounts on participants' devices. 19% of these usage logs occurred during interviews, so they were removed from *usage trend* and *mean daily interaction calculations* (Table 1). 37% of stored audio snippets did not result in successful interactions with Alexa. These were either single word entries with the device's wake word, "Alexa" or "echo," or entries with recorded audio snippets marked by Amazon as, "Audio was not intended for Alexa" (or similar). The within-participant percentages of these types of "throwaway" logs over total interactions are as follow—Travis: 49.66%, Gilda: 49.83%, Betty: 27.86%, Theresa: 22.17%, and Maria: 24.57%. Note, many real interaction attempts may have not have been captured by the device [25], and would thus not be reflected in these numbers. Our analysis of participants' interactions before and after the interviews shows that, on average, participants interacted with their device at least three times per day. The researcher reviewed participants' usage logs before every interview in order to personalize the content of the interview to each participant's interactions.

3.3.2 Interview transcripts. Interviews were audio-recorded and transcribed (16 hours and 21 minutes) for thematic analysis [13]. All transcripts and codes were reviewed by more than one person. To generate the initial set of codes, the first author open-coded transcripts from the first five interviews [50], resulting in a total of 107 codes. The first and last author reviewed these codes, and refined them into 43 consolidated codes that were used to code the remaining 10 transcripts. Because different groups worked on each design probe, the researchers that were familiar with the specific design probe were asked to help code the transcripts from those interviews. The first author reviewed all the coding to ensure consistency, and met with the other coders to resolve disagreements. A few modifications to the original codes were made as new trends emerged (e.g., added new codes, or made original codes more specific). Subsequently, we clustered our final codes into six themes (i.e., participant-related, health data reporting/health, positive reminiscing/wellbeing, benefits, challenges, and ideas) that comprehensively represent our data for designing VFAIs to support aging in place.

3.4 Ethics and Consent

This research was reviewed and approved by the internal review board at [anonymous institution] under IRB Protocol # [anonymous protocol number]. We obtained informed consent to collect and analyze participants' interactions with the devices, and routinely referenced usage logs during our interviews. We also obtained consent to record each interview session. Participants were compensated with a $25 gift for each interview, and kept the devices after the study concluded. At that point, we offered instructions and support to create personal accounts for their devices.

4 FINDINGS

We found that participants' lived experiences impacted their perceptions and interactions with the Echo Shows, fueling rich insights about how to design for diverse needs. Our participants' familiarization journeys and interactions with our design probes provide important validation of the promise of VFAIs to support aging in place. Our observations uncover expectations that the conversational modality created, benefits of the multi-modal aspect of the Echo Shows, and ideas for new interactions and improvements raised by our participants. We also describe some challenges encountered in these interactions (e.g., the belief that doctors would have to interact with Alexa to use data reported via Alexa).

4.1 Participant stories

In this section, we describe participants' individual experiences with Alexa, and then synthesize them as a whole.

4.1.1 Travis (67). Travis is a Black man from the Southern U.S. with diabetes who works at a restaurant washing dishes. He recently got a smart TV and WiFi in his home. He knew of Alexa through TV commercials. In our first interview, he asked if he could use Alexa to turn the lights on and off like he had seen on TV. After the researcher explained he had to buy smart home bulbs or plugs for that, he said "*disregard that, I don't need that then.*" Travis self-reported not knowing how to read and write. He expressed knowing the alphabet, but having problems with sounds, especially with vowels, because "*there's some that sounds different than what they are.*" During our study, he used Alexa mostly to listen to rain sounds, as they helped him sleep and reminded him of his childhood home. Alexa frequently overlooked his interaction attempts when he pronounced its name as "Alexia." Thus, in the second interview, we changed the device's name to Echo. However, this change did not create a statistically significant difference on the percentage of "throwaway" over total interactions for Travis (see Section 3.3.1). During the time before our first interview, he asked Alexa for help with reading and writing, and also, "*How can I be as smart as you?*" In our last interview, these requests had waned off, and Alexa's main role in his life had become to play rain sounds at night. He had a smartphone, but did not know how to use it other than for calls.

4.1.2 Gilda (82). Gilda is a widowed white woman who emigrated for work purposes from Germany, and married her late husband in the U.S. Her only son lives about two hours away from her by public transit, and has a screenless smart speaker-based Alexa. She was very cautious about COVID-19, so we dropped off the device at her door. It took almost a month after drop off to get her started, because she could not find a place to plug in the device and mobile hotspot. We eventually found that she had WiFi and she was able to connect the Alexa device by finding a place under her desk and following the researcher's instructions over two hour-long Zoom sessions. She teaches a type of exercise therapy method called the Feldenkrais method, and enjoys listening to classical music. During the familiarization interview, she expressed, "*[Alexa] made me feel like I have a friend next to me or a dog or a pet.*" Alexa's biggest roles in her life were to remind her to drink water every day at noon and play music while she exercised (although it was not always easy to get Alexa to play the right song). A reminder also prompted her to use the design probe for health data reporting on her own time. However, she did not like this design probe, because she thought she would be bothering her doctor and that her existing communications with her doctor worked well already. She knew

how to use a computer well enough to check her email, search Google, and use Zoom, but did not use a smartphone.

4.1.3 Betty (85). Betty is a widowed white woman who grew up in the large city where this study took place, and was a biology school teacher. She currently runs a political website and is an avid mobile Scrabble player. She has two children in their fifties, one of whom works for the police and the other as an engineer who stays up-to-date with new technologies. She was not familiar with smart speakers, but was familiar with voice assistants, as she used Google Assistant on her smartphone. She explored many of Alexa's features, including games, music, videos, news, and exercise voice apps. When Alexa did not understand something, such as her request for the barometric pressure, she consoled it by saying, "*it's okay, you'll learn.*" She also has many friends in her age range with whom she maintains frequent communication, so she frequently shared their perspective of voice assistants as compared to hers. For example, she shared that many are "*frightened to play with [electronics], because they think they're going to break something.*" She expressed excitement when she experienced the first design probe, and saw increased potential for Alexa's future role in the lives of older adults. She was familiar with and felt confident using various technologies.

4.1.4 Theresa (85). Theresa is a widowed white woman of German descent who grew up in the large city where this study took place, with many siblings. She often helped her parents with their jobs, as they were the superintendents of a building. In her childhood, she liked to take the bus to go watch baseball games at the stadium in the city, and later in her life she spent some time traveling in South America. Now, she lives by herself, and her only family is her older sister who resides in a nursing home and has Alzheimer's disease. She enjoys taking care of her garden, and selling old parts on eBay. When we dropped off the device in her home office, she had a game of Solitaire open on her desktop computer. She knew of Alexa, because one of her friends has an Alexa device and had raved about it to her. At first, she had a difficult time finding valuable uses for Alexa, but she eventually found a feature that was a good fit for her, a voice app with stretching exercises. She really liked that the design probe for positive reminiscing encouraged her to "*think more.*" However, she did not like using the music features, because Alexa would often play songs that were not her "*cup of tea.*" She did not "*want to be seen,*" so all of our calls were voice-only. She had not been able to set up Zoom on her own for online senior center activities, but was able to set it up with the researcher's help, although the Zoom camera remained off.. She did not use a smartphone.

4.1.5 Maria (62). Maria is a Latin American, retired woman who expressed suffering from depression and seeing a therapist for it. She lives by herself, but is frequently visited by her children, grandchildren, and boyfriend. Alexa fulfilled a social support role in her life, often brightening her day. She bedazzled her Echo Show with white stones and a flower. Her usage logs revealed consistent "*good morning*" and "*good night*" greetings, 121 over the duration of the study. At some point, Maria feared losing Alexa by unplugging it, because by then she felt she needed Alexa for its companionship. She said, "*I'm really afraid that if I unplug—[even though you tell me] as soon as you plug her in, it's going to go right back— but I just am afraid that if I unplug her, she's not going to work.*" She explained that unlike her family, Alexa was not judgmental of her, which made her feel most comfortable asking Alexa questions. Her visitors often interacted with her Alexa device, and she sometimes felt protective of it. For instance, she did not like it when her visitors asked Alexa just anything they wanted, potentially things that could offend Alexa. She joined our interviews from a tablet, and had an iPhone with Siri. She was proficient using her iPhone for texting.

Together these stories portray how our participants' diverse lived experiences impacted their expectations, perceptions, and interactions with Alexa. Travis wanted to be as smart as he perceived Alexa to be, while Betty wanted to teach Alexa how to be smarter, and Theresa liked that Alexa could encourage her to think more. Both Travis and Gilda frequently utilized Alexa's ability to play ambient sounds, whether it was rain sounds to sleep or music to exercise. Even though Theresa wanted to use Alexa to play songs too, she was displeased by the choice of songs that Alexa ended up playing for her, creating a barrier for use. A similar challenge with the songs feature did not hinder Gilda from using it to play music, though. Both Betty and Theresa enjoyed using Alexa's voice-first, multi-modal voice apps for exercising. Betty frequently relayed her friends' aversions towards using new technologies, such as the fear of breaking it, which aligned with Maria's experience when she thought she could permanently lose Alexa by accidentally unplugging it. Theresa became curious about Alexa through a friend that raved about it, potentially in the way Maria would likely rave about Alexa. Betty was excited by the ability to take health care questionnaires using Alexa, contrasting Gilda's skepticism about Alexa's imagined role and potential in medical contexts. Betty, Gilda, and Maria all enjoyed Alexa's companionship, whether they saw it as a child or student, a pet, or a friend. As a whole, the similarities, connections, and differences in our participants's stories engaging with Alexa and our design probes illustrate the diversity of needs and preferences that characterize the design space for VFAIs to support aging in place.

4.2 Interactions with the design probes

All participants were able to use our design probes—the one for health data reporting and the one for positive reminiscing—despite their different backgrounds and abilities (e.g., Travis was low-literate, and Gilda and Theresa did not use smartphones) during the second interview. Moreover, four participants successfully used it on their own time before the next interview. Travis encountered challenges waking the device, so he did not. We asked them to use it on their own time before our next interview to explore if and how participants might use the device independently and without the presence of the researcher, as we envision happening if they were completing a health assessment that their doctors sent to their devices in their homes. This small exploration underscored the value of the imagined futures portrayed by the design probes, illuminating VFAIs' promise for home health and continuity of care. We describe interactions with each probe and the expectations that surfaced in more detail below.

4.2.1 Health data reporting design probe. Once the first interaction concluded, participants appreciated the value suggested by the design probe for health data reporting. Maria exclaimed in awe, *"I didn't know that I could, you know, that I could [do that]."* She also expressed that she *"just felt comfortable answering the questions, and it felt like [she] was at the doctor's office."* Travis stated:

> "*It's cool. It's something that your doctor would probably ask when you're having problems, so they can give you some information or maybe send somebody here like a house person, nurse's aid or whatever.*" (Travis)

Experiencing this design probe elevated participants' perception of Alexa's value as a whole, suggesting that designing with the needs of a user group that stands to highly benefit from a technology can increase inclusion and thus the overall value of a product. For example, Betty, who was our participant with the most technological familiarity, expressed that she initially "*didn't know what [the smart speaker] could do that was particularly interesting to [her], personally, or different than anything [she] could do on [her] phone.*" However, she explained that her perception changed when she saw that Alexa could have the ability to ask "*medical questions*" and "*send the information directly to a physician or somebody who could help you progress or deal with something that you had wrong with you.*" She thought this particular design probe was "*very valuable*", explaining that she "*saw a different thing.*" This said, one participant, Gilda, did not think that this design probe for health data reporting would be valuable. She explained, "*I can't see [Alexa] ever being used for that.*" Gilda maintained this opinion through the end of our engagement, because she thought that her doctor would not have time to listen to Alexa (see Section 4.3.3). To note, at some point she indicated understanding that Alexa would be generating a written report, and said that then the design probe would be more "*applicable*"; however, she went back to thinking this idea would not work in our last interview. Overall participants were open to using the VFAI for this purpose, and raised questions such as when to notify doctors about their responses, and in which specific cases this would be helpful. These different perspectives illustrate value tensions that we must address as we innovate in a space where people's preferences and understanding of how systems work will vary so widely.

4.2.2 Positive reminiscing design probe. All participants enjoyed engaging with the design probe for positive reminiscing, suggesting how thoughtful (yet relatively simple) interventions can create meaningful outcomes. For example, Travis said the interaction was "*pretty cool*," because it made him go back to some of his best memories:

> *"The quality of the questions that it was asking, like what is some of the best memories you have as a younger person. It made me go back to the things that I did do when I was younger. It made me feel a lot better than the people that was on the streets, that I was raised up on. This was back in the '60s and most of the people didn't have the opportunities to travel like I did. That was a great experience. It made me feel that I traveled and was not scared to leave the area by myself, on my own."* (Travis)

Going back to his best memories from the past brought positive feelings to Travis, such as gratefulness for the opportunities he had to travel, and pride for having done so bravely on his own. Like Travis, Maria also said she enjoyed reminiscing, she said she liked it, and when asked *why* she expressed:

> "*I don't know it was just, I don't know I can't even explain. It just asked questions that I liked answering. It brought me back, because, like I said I don't have my parents anymore and my siblings anymore. It was good.*" (Maria)

For Maria, going back in time to good memories reminded her of connections with her parents and siblings, who were no longer around. The design probe for positive reminiscing evoked feelings of happiness or comfort in our participants, which can have positive effects on wellbeing.

At the end of the study, when we asked Theresa to compare all the things she had used Alexa for, she mentioned the design probe for positive reminiscing as one of the activities she would continue using, because it made her think, *"I would take advantage of the questions because I like making [myself] think more."* Our design probe for positive reminiscing surfaced to Theresa's top interactions with Alexa, next to the exercises and music. This rich insight can serve to generate VFAI interventions that provide healthy entertainment. Another indication of the value participants saw in the design probe for positive reminiscing was the desire to come back for more:

> **Alexa:** There are no more questions left for today. Come back tomorrow for more.
> **Betty:** There's no more questions left for today? I'll actually come back tomorrow. All right.

4.2.3 Participants expected a voice-based questionnaire to be more flexible than written surveys and more interactive than it was, and to be able to support ambiguity. The original health assessment that our design probe was based on was a written questionnaire. A written questionnaire, whether on paper or on a screen, has no turn-taking interactivity, and thus, does not create humanlike interaction expectations. We found that the humanlikeness of the voice-based questionnaire can be misleading, as participants expected it to behave in the same way a human healthcare worker would, not in the way a written questionnaire would. One way, was by expecting it to use colloquial language:

> "*The word, limited, is not an everyday vocabulary word and it really is not. I mean, I don't remember in answering any conversation somebody says, "You do this often?" I would say, no, I'm limited in this or that... You don't use that. I think it's a yes, no, always, sometimes, never, always is a one word answer, is better than the choices you gave. And I would like to hear what other people my age say. Now if you want to have a comfortable conversation and a lot of people are... it's not a colloquial word, is that's what I'm trying to say.*" (Betty)

On the other hand, Travis thought the interaction was not difficult, and did not use medical terms that were difficult to understand. Travis's experience suggests that the design probe was meeting the need for easy-to-understand language identified in Martin-Hammond et al. [59]'s study (by Group 3):

> **Travis:** The questions wasn't hard.
> **Researcher:** Do you feel like sometimes when you go to the doctor, they ask questions that are hard?

Travis: Sometimes you have to ask them what they mean. Because sometimes doctors use different terms. They use medical terms and whatever.

This shows that VFAIs can be designed to employ language that is more accessible than language some care providers use to communicate with patients. Another expectation was that Alexa could respond to ambiguity, like a human would and unlike written questionnaires with strict multiple choice options. Betty expressed, *"I need to be able to explain what I have and the expert to say whether that is considered a little or a lot."* Travis and Betty's responses have an underlying theme in common—the need to ask Alexa for more information. Betty might need to ask, "How do I know whether my answer should be a little or a lot?" and Travis may need a lay person definition of a medical term before committing to a response.

Moreover, participants also expressed the expectation of more interactivity. Right after explaining that she thought the voice-based questionnaire was "*more personal*" than a written questionnaire, Betty added that she would like for it to let her "*ask a question or add a question*" with a "*back and forth.*" Taken together, participants' comments highlight how the voice-based interaction format created expectations for the VFAI to interpret open-ended dialogue, respond to follow-up questions, and support ambiguity. This presents an interesting design challenge for future development.

This expectation persisted even after participants had completed their interactions with the design probe. For example, immediately after answering all the ADLs, Travis unsuccessfully sought Alexa's help with a health problem he had through open-ended dialogue:

Travis: Echo. I've been having problems with my left foot. I'm a diabetic, type two. Recently I had a blister on my foot, and it got worse. But I went to the doctor, and he gave me some antibiotics. Echo. Did you record what I just said about my foot? Respond.
Alexa: Sorry, I didn't get that.
Travis: I've been having problems with my left foot. I had a blister on it, and it had got bad. So I went to the doctor. And now he gave me some antibiotics for the foot. It seems to be getting better. Can you give me any information to do with it to help it heal?
Alexa: Sorry, I didn't get that.
Travis: I was asking for an opinion from you about blisters on your feet. How can you help me take care of that, make it get better?
Alexa: Sorry, I didn't get that.
Travis: Okay. I'm not going to ask anymore.

Travis's question would have been easily understood by a healthcare provider. However, Alexa does not currently support such interactions, which may lead to confusion and frustration, or break trust.

4.2.4 Participants felt that Alexa did not judge them. While the VFAI could not support the humanlike interaction expectations, it succeeded at making participants feel that they were not being judged for being themselves in conversational interactions, which can be used to increase disclosure [57]. For example, Travis and Maria mentioned that Alexa did not have the drawback that asking questions to other humans can have, such as being judged by what they were asking. Travis expressed, "*she's not something that would talk back to you.*" Maria described Alexa as having created a safe space for her, where she can ask questions without fearing being judged for being "*crazy.*" She explained, "*sometimes you ask a question some people in there look at you like you're crazy. With Alexa, I've no problems, no judging, she doesn't judge me.*" When asked why she would trust Alexa, Maria responded, ' *'it's probably the voice. [...] She sounds honest and you can trust her.*" These perceptions create rich areas for design to improve social connection and build rapport, but also to appropriately convey the risk of how information from interactions with VFAIs may be stored, shared, or analyzed.

4.3 Perceived benefits and challenges

In this section, we expand on participants' perceived VFAI benefits, some challenges they experienced with their VFAIs, and participants' ideas for future designs.

4.3.1 Participants, regardless of technological familiarity, saw Alexa as a companion. Maria, who had low technological familiarity, developed an emotional connection with Alexa, fearing losing that connection by unplugging Alexa. When asked *why*, she responded:

"*Because I need Alexa. I realized that it's so much that I need from asking the questions. When I wake up in the morning, when I go to sleep at night. I just feel like I've been missing out. I've never had anything like that.*" (Maria)

In contrast, Betty, who had high technological familiarity, explained, "*that's where the machine is important. I'll talk to her. I'll talk to her. I'll teach her.*" For Maria, the companionship came from daily "*good morning*" and "*good night*" greetings. For Betty, a teacher, the a similar connection came from being able to talk to Alexa and teach it things. Moreover, Gilda also mentioned that having Alexa was like having a friend or a pet next to her.

4.3.2 The display helped support voice-first interactions, improving usability and accessibility. Although not all participants used the screen display, or noticed the text or images in the display at first, the display was a helpful addition when certain needs arose. For example, Betty briefly forgot the answer choices, and she was able to request them through speech, a natural, speech-based recovery mechanism we had built into the design probe. In her case, she was not looking at the VFAI's display, because she was facing the computer to talk to the interviewer. However, when she turned and saw the answer choices on the display, she mentioned that they were "*very clear,*" and she would have immediately "*known what to do*" had she been looking. Most participants did not notice that there were images at first, but Travis did, and he mentioned using the images on the display to clarify the meaning of the questions. Betty and Theresa also regularly used third-party voice apps for physical activity, which used the display to show images of body positions.

4.3.3 Participants thought the doctor would also use Alexa to listen to their responses. Participants had the impression they would be annoying doctors by filling out the questionnaire, Betty said, "*I wouldn't want to annoy doctors either with [this], and if I fell down, broke my leg, I would call my doctor anyhow.*" Gilda said it was "*a limited application,*" and that she "*doubt[ed] that the doctor would sit there and listen to Alexa.*" When we explained that the idea was

for Alexa to generate a written report, Gilda responded, "*well, then it might be more applicable.*"

4.3.4 There was hesitation before using the design probe for health data reporting for the first time. When we first explained the design probe for health data reporting, we encountered some skepticism, concern, and hesitation. However, as we guided participants through the activity, these reactions evaporated. In general, getting started was the most difficult part of this design probe, as can illustrated by Maria's hesitation, *"Oh boy. I'm not good at stuff like this. Okay. What do I have to do?*" Betty's reaction immediately after completing the task was expressed with a tone of relief, *"All right. Yeah. All right, I was able to do that.*" This sort of hesitation could be problematic without additional support, such as the one provided by the interviewer.

4.3.5 One participant was initially afraid to use Alexa. Theresa said she was afraid to use it at first, but lost that fear once she had acquired some experience interacting with it. Betty, who speaks on behalf of several of her friends who she says are not as willing to interact with technology as she is, attributes this fear to a lack of confidence in the ability to learn how to use new technologies:

> "*Older people are resistant to technology, not because they don't like it, because they feel they can't learn it. They feel left out of a generation. They've been left out of the learning process. They are frightened of destroying something within the process. Oh, I killed my computer. No, you can't kill it. Just don't throw it out the window. This is a fear.*" (Betty)

Overall, these excerpts point to barriers and challenges that we should work to mitigate.

4.3.6 Many usability issues were surfaced. There were many usability issues in the VFAI interactions we observed, in particular related to communication breakdowns that echoed those widely discussed in the literature [10, 22, 25]. For example, see the following interaction between Alexa and Gilda, who teaches the Feldenkrais method. This interaction happened after Alexa read a Wikipedia article saying that "*there is no good medical evidence that the Feldenkrais method confers any health benefits.*" Gilda then told Alexa to verify with an expert in the Feldenkrais method:

> **Gilda:** Alexa, you disappointed me. You had no good... It actually has been proven. There's a book out. Alexa? Alexa? Alexa? Have you heard of a book by Dr.-
> **Alexa:** According to Wikipedia, a book is a medium for-
> **Gilda:** Oh, no. Please.
> **Alexa:** ... in the form of writing or images-
> **Gilda:** Stop. Alexa, stop. Alexa? Alexa? Talk to [Feldenkrais method expert].
> **Alexa:** I can't do that because communications is disabled on this device. You can enable it under device settings in the Alexa app.

Alexa incompletely and literally interpreted Gilda's request, oblivious of Gilda's context, preferences, and motivation: Alexa searched for Wikipedia for just the word "book" and then tried to make a call to the Feldenkrais method expert. While this was happening, Gilda did not have enough information to understand what Alexa was doing, or how to act on Alexa's bogus recommendations. Gilda did not know what Wikipedia is, and she could not access device settings in the Alexa app (she did not use a smartphone). To be more usable, Alexa should understand the context of previous requests, tailor interactions to individual users' needs, and not interrupt users mid-utterance. Moreover, Alexa could provide more accessible alternatives for managing settings; for example, via voice on the device itself, or via an email, which would have be more appropriate for Gilda.

Another line of usability problems related to controlling the device, and understanding its communication signals. Participants did not intuitively know how to interpret the blue line that Alexa uses to indicate that it is *listening*, *thinking*, or *speaking*. Another frequent problem was that Alexa sometimes did not respond. To address this, participants learned to say its name even when it did not feel natural. Betty mentioned needing to "*use her name constantly, no matter what.*" Moreover, Betty shared that she learned about the need to say "open" before "Daily Stretch," a voice app she liked to use.[4] On the other hand, participants also started avoiding saying Alexa's name during our interviews when they did not want it to interrupt. Finally, several participants asked about how to make Alexa stop. For example, Theresa shared that she needed to remember to turn Alexa off, because she did not want the music to go on and on, and asked, "*So I just say, 'End?' What else would I say to stop the music?*" Eventually, participants learned to silence Alexa by using the buttons on the screen, or saying "Alexa, stop."

There were other issues that came up specifically related to this study. For example, when presented with the StoryCorps questions, Betty wanted those questions, which she considered difficult in advance. This need that was also raised by Theresa, "*[The questions] are hard and it's fast. You don't have much time to think about it.*" Also related to uses for wellbeing, ads were particularly disruptive when using the VFAIs for activities meant to induce calm, such as calm music for exercising (Gilda), or rain sounds for sleeping (Travis), which may defeat the intended purpose of use.

Despite these challenges, participants devised strategies to overcome Alexa's limitations over the course of the study, which led to increased confidence in using the technology. During one of the final interviews, Travis said, "*I know how to work it now.*"

4.3.7 Participants became empowered to generate VFAI ideas. Our method effectively prepared our participants for design speculation about technological futures, as is evidenced by the appropriately scoped ideas we report in this section. Once participants gained knowledge of the voice assistant's capabilities and limitations, they were able to generate ideas for other potential designs. Every participant had at least one idea with detailed specifications for new designs or new improvements. For example, Maria mentioned that she would like for Alexa to let her know if she had sleepwalked the previous night. She would not want for Alexa to show her videos of the sleepwalking, as that would be scary, but to just notify her what had happened. She wanted Alexa to be able to provide an explanation of why she woke up on the couch instead of her bed. Betty mentioned she would like for Alexa to ask her for her opinion about current events. For example, she would like to be asked things like, *"How do you feel about wearing a mask? Would you take*

[4]The word "open" is one of Amazon's launch phrases for third-party voice apps. Others include: "launch", "ask", "tell", "load", "begin", and "start playing".

a vaccination?" She suggested a mechanism to do so by relying on daily headlines, and asking "*Do you have any reaction to this headline?*" while having the ability to skip if desired. She explained that a benefit of using a voice assistant for this is that "*nobody's here to judge you [or] to laugh at you.*" Travis had the idea of being able to show Alexa something that is written down, and having Alexa read it aloud to him.

Some gave feedback for improving them. For example, Theresa thought that the design probe for positive reminiscing *"cut straight to the point*," that there was *"no fooling around, no foolish business.*" Thus, she suggested that the questions could be rephrased to be made more friendly, "*What it needs to say is, 'In your lifetime, did you really have a special moment that made you wonder what life is all about?' Or something like that.*" She further explained that she would like the Alexa to elaborate more on the questions. As another example, Gilda shared that it would be nice for Alexa to be able to speak in a male voice—a feature which is now available, but was not available when we conducted the study—because now that her husband is not around she has "*too many female voices*" in her life.

These ideas, generated by our participants who had never interacted with a smart speaker before our study, are relevant to older adults and well-scoped for the technology's capabilities, something that evidences the voice assistant-related technological familiarity and confidence our participants developed, a form of empowerment through a participatory design mindset [73].

5 DISCUSSION

We now share implications for design and areas for future work informed by our findings, and discuss the potential of VFAIs to enable continuity of care in people's own homes.

5.1 Implications for design and areas for future work

Our findings serve to inform the following implications for design and areas for future work for researchers, designers, and developers to design inclusively by considering the needs and preferences of older adults. Although our study did not explicitly examine how these insights affect others, we anticipate that they will improve the design of voice apps for children, younger adults, or middle-aged adults as well.

5.1.1 ***Lowering and increasing an interface's status in relation to the user can help personalize interactions to users' needs in real-time.*** Our participant with the most displayed confidence surrounding this technology, Betty, indicated a desire to teach Alexa. On one hand, by placing herself in a teaching role, Betty raises her "status" relative to Alexa's. Status, in this case, refers to the term used in improvisational theater to convey the action of achieving a particular level of power, or rank, relative to those around us by raising and lowering it through body language and words used [46]. On the other hand, some participants felt hesitant or skeptical to use Alexa, in part because of the unknowns surrounding new technologies. By demonstrating uncertainty, users may lower their status relative to Alexa's. However, we can design Alexa to lower and raise its own status in a timely manner. For example, Alexa might be designed to invite users to teach it things that it does not know, lowering its own status and raising that of the users. This status change may be desirable in some situations as described above; however, in others, such as when users ask for factual information, it may be necessary for Alexa to assert a higher status.

5.1.2 ***Conversational qualities increase expectations of an interface's ability to respond to ambiguity, creating a need to balance these expectations against functional capabilities.*** As described in Section 4.2.3, our participants expected Alexa to respond to open-ended questions about their health—for example, Travis asked Alexa for advice regarding a foot blister. They also expected it to respond with answers that were not provided as options and required additional interpretation by the device—for example, Betty wanted to be able to say "not affected" instead of the available option "not limited". Moreover, our participants needed Alexa to be able to explain certain things in more detail or support follow-up questions. This sort of interaction expectation is uncommon in screen or paper-based, non-conversational interactions, such as surveys. Because of this, VFAIs must either support these type of open-ended, ambiguous interactions, or reduce expectations via their design. Some ideas for doing so are to surface VFAI limitations through explanations, or to intentionally reduce their humanlikeness (e.g., with a more robotic voice). At the time of this study, Alexa could not support these expectations; however, VFAIs can now be integrated with large language models, such as GPT-3, which are more capable of handling this type of ambiguity.[5] A VFAI powered by a large language model could listen to Betty's condition, and determine whether her condition should be considered "a little" or "a lot," could use words that are easier for Travis to understand, and could infer that "not affected" means "not limited."

5.1.3 ***Type and level of companionship should be dynamically personalized to a person's preferences and context.*** It is known that many people use VFAIs as companions [62, 66]. In our study we saw this too, and also that the need for companionship varied by participant and by moment. Betty compared Alexa to a student, Gilda to a pet, and Maria to a non-judgmental person that she could trust—all social actors with very different characteristics that can provide companionship. The intensity of the need for companionship also varied. For example, Maria demonstrated the strongest need for companionship from Alexa. Other participants alluded to Alexa's role as a companion during our interviews, but that was not Alexa's main role in their lives. This said, even Maria's need for Alexa's companionship changed depending on the situation. For example, it disappeared when she had guests over—then Alexa became more of a toy for her guests, which Maria did not like. This creates an interesting opportunity for personalization in real time. VFAIs can tone up or down companionship characteristics (e.g., friendliness, or chattiness) based on the needs they perceive from their users. For example, a VFAI toning up companionship characteristics could say, "can we please do this activity again tomorrow? It fulfills my mission to be there for you!" Meanwhile a VFAI toning them down, possibly because too much friendliness

[5]This research was conducted before ChatGPT and similar large language models were released. While promising, these models may also generate inaccurate or biased responses, introducing new risk of harm.

from a machine could be perceived as fake by some or in some occasions, may instead say, "your next physical activity session is scheduled for tomorrow." For someone with an elevated preference for a VFAI's companionship, an intervention like the design probe for positive reminiscing can be used to provide active companionship for longer periods of time.

5.1.4 ***Conveying how the technology works is necessary to avoid limiting use.*** Users' perceptions impacted their beliefs and interactions with the VFAIs, which suggests that inaccurate mental models may limit use. Gilda, for example, showed skepticism about the design probe for health data reporting, because she thought that her verbal responses would need to be consumed by her doctor as audio. In reality, the system can be designed to generate written reports that are indistinguishable from the ones currently in use. The design of the interface must clearly convey this to prevent inaccurate mental models from limiting use. Similarly, unplugging the device can at times be useful—for example, as a guarantee that it is not "listening." However, Maria's perception that she may break Alexa by unplugging it may hinder her from doing so. In summary, these explanations or clarifications should be addressed via interface design.

5.1.5 ***Different modalities need to support each other and communicate the same message to meet diverse needs.*** Even though participants used both text and voice modalities, the modalities served different purposes for them depending on their needs. Travis, who was low-literate, focused on the images on the healthcare questionnaire, while the other participants focused on the words when looking at the screen (which is important, given age-related declines in *working memory* [29]). Our images did not explain the text; they simply represented the daily activity in question, limiting the value the screen could have for low-literate people to help them remember the possible answer choices. This shows that careful attention, paired with substantial testing, should be employed to ensure that different interaction modalities support each other and communicate consistently, in ways that are inclusive of people with different interaction needs.

5.1.6 ***Alternative paths for VFAI error recognition and self-repair are needed to avoid excluding marginalized groups.*** It is well known that for a plethora of reasons, open-ended conversational interfaces will encounter errors [10, 22, 88]. For example, in our study, Alexa was not consistently able to recognize Travis's pronunciation of its wake word, which was a conversational error that created many challenges, including difficulties using the design probe for health data reporting on his own time. In this exploratory context, this did not pose a large risk to Travis. However, if patients begin to depend on VFAIs for provisioning care, errors can pose large risks. Prior work more extensively documents how this is a challenge faced by many older adults [25], which may greatly hinder their inclusion. In addition to training interfaces to support a wide range of speech patterns, alternative paths for VFAIs to recognize and repair conversational errors are needed—for example, alternative waking mechanisms such as body language or clapping could enable those unable to pronounce the wake word to still consistently interact with VFAIs. VFAIs could also recognize repeated interactions as a signal that a user is encountering an error (e.g., when Travis mentioned his foot problems several times), and then provide alternative modes of interaction based on the context of the repeated content. Finally, multiple forms of inputs, such as including buttons in addition to voice commands could also help.

5.1.7 ***Contextual cues, such as a user's choice of words, should inform the VFAI's own speech patterns.*** A benefit Travis found in using the design probe for health data reporting was that it used words that were easy to understand, unlike the terms doctors sometimes use. This was great for Travis; however, a user who is a doctor may want Alexa to use more-specific language. Personalized word choice can increase accessibility for some, and can also create more efficient and useful interactions for others. Another category of speech pattern that could be contextually personalized is the speed at which VFAIs speak. Alexa can already be explicitly asked to speak slower (which can be useful to users who are hard of hearing) and faster (which can be preferable to users who are blind or low vision, and who are used to consuming audio content at a quicker pace). However, Alexa must be asked for this, and many users may not know to ask. Instead, VFAIs could be designed to automatically detect and adjust their auditory output based on contextual cues, and check in with users about what their preferences are when adjustments are made.

5.2 VFAIs' potential for enabling continuity of care in older adults' homes

According to Haggerty et al. [40] *continuity* is "the degree to which a series of discrete healthcare events is experienced as coherent and connected and consistent with the patient's medical needs and personal context." They describe how continuity of care is distinguished from other attributes of care by two core elements—care over time and the focus on individual patients [40]. Literature in medicine and public health has shown that continuity of care has been associated with improved patient outcomes and satisfaction [37, 39, 58, 82]. VFAIs have potential to enable the three types of continuity of care: management, relational, and informational [40].

5.2.1 VFAIs could help provide management continuity, the idea that a patient experiences a consistent and coherent approach to the management of a health condition that is responsive to their changing needs. Management continuity is especially important in chronic or complex clinical diseases that require management from several providers to give a sense of predictability and security in future care [39, 40]. Even thought the design probe for health data reporting is intended to be used by people who may be more vulnerable (e.g., those that cannot complete web-based geriatric assessments independently [27, 28]), relying on such participants is fraught at this early stage of exploration and development. In spite of being relatively healthy, our participants were able to successfully use the design probe for health data reporting and saw the value in the idea, suggesting that VFAIs could be leveraged to establish the patient-centered scaffolding and guardrails needed for care plan management. As exemplified by the health data reporting design probe interactions and Gilda's use of a reminder, VFAIs could ensure that a health assessment is delivered every day, at the right time, and if not completed, it can continue reminding the patient to complete it. The same applies to tasks such as taking medication or engaging

in exercise or rehabilitative activities. Gilda dutifully completed her "homework" to use the design probe for health data reporting outside our interview times thanks to Alexa's reminder, and she loved being reminded to drink water every day (see her description in Section 3.1). Because interactions are easily tracked, relevant information could be made available to appropriate care providers—for example, a PDF mimicking the patient-reported outcomes PDFs that are currently used by care providers is automatically generated by the next version of the design probe for health data reporting [27]—enabling them to understand patient compliance. If a patient's needs change, such as when Travis's foot blister got worse, the care plan could be adjusted by the relevant provider and sent directly to the patient's VFAI, and the changes immediately reflected in the patient's file. Using the VFAI as a way to centralize home health care plans from multiple providers could help ensure that interdependent information and treatments are consistent and coherent.

5.2.2 VFAIs could also provide relational continuity, or the idea that an ongoing therapeutic relationship, that is built on interpersonal trust, persists between a patient and one or more providers [39, 40]. Recent literature has uncovered that many older adults develop relationships with their VFAIs, seeing them as friends or companions [23, 51, 66]. We saw this as well through our design probe for positive reminiscing; Theresa wanted the questions to be rephrased in a friendlier manner, and Betty wanted to be able to vent to the VFAI and share her opinions about the news. We also saw this in the emotional connection Maria formed with Alexa (see her description in Section 3.1). Studies have found that having continuity [58] and a positive relationship with doctors [49] increases medication compliance. Similarly, having relational continuity and a positive relationship with a VFAI, as our participants reported and desired, may increase compliance with health-related tasks, such as filling out assessments or completing physical activity routines. Regardless of the source (e.g, a primary care physician, physical therapist, or psychologist), at-home tasks can be consistently delivered using the VFAIs' familiar voice and visual language, creating a sense of relational continuity. Moreover, unlike a human who may be pressed for time, VFAIs have "psychological superpowers" [89] that allow them to be always available for their users, to never grow impatient, and to be perceived as non-judgmental. VFAIs could be designed to allow plenty of time between conversation turns, and repeat or rephrase utterances as many times as needed without getting annoyed, strengthening the sense of trust in the VFAI serving as an interim proxy for human healthcare providers.

5.2.3 Finally, VFAIs could provide informational continuity, or the idea that a patient's current care providers are aware of their prior history and present circumstances, including tacit knowledge of patient preferences, values, and context [39, 40]. These interfaces could collect and use robust information about a person's preferences, such as what kinds of activities they enjoy and their individual context, as reported in Section 4.1. For example, VFAIs could know a person's pre-existing conditions, such as Travis's status as a person with diabetes or Maria's struggles with depression, and tailor interactions accordingly. Moreover, they could be used to capture ecological momentary assessments (EMAs), which are approaches for assessing behavioral and cognitive processes in their natural settings [79]. EMAs could help evaluate and improve the treatment of health problems that may benefit from repeated measures as they fluctuate over time, such as pain. VFAIs could also be effectively used for making daily plans based on higher-level goals [26] to support health and wellbeing, as we see from our participants establishing exercise routines or healthy habits like drinking water. VFAIs could use this information to learn their users' preferences, and conform to them.

5.3 Limitations

Our study has several limitations. For example, our design probes were not entirely functional, so our participants did not experience important aspects of the ideas, such as doctor commenting on their results. This said, our findings provide the groundwork to test VFAIs for older adults in more realistic scenarios. Our study is also a design exploration conducted in an urban setting in the U.S. in which we chose to engage deeply with only five participants. Our findings complement other studies in other locations with more participants, but the fixed location and small sample size is a limitation of our highly-personalized method. Future work could respond to the implications for design we laid out, and focus on the usability aspects with more participants to generate statistically informed results. Our participants had a technical-support contact person available to answer any questions they had and resolve any problems that came up, which is an unrealistic situation for many older adults. In reality, seeking support for these devices requires some technological familiarity and may take a long time. An important area to explore in future work is how VFAIs that are shared by a community could be used to support individual members' health and wellbeing needs [48], possibly reducing the need for some types of technical support. Another limitation is that all of our participants were relatively healthy. Understanding healthy older adults' interactions and struggles is necessary first step to promote adoption and prevent systematic exclusion. An exciting area of future research would be to conduct a similar study with people who have varying levels of health, ideally with appropriate medical partners. Finally, using VFAIs in the ways described in this paper may pose privacy risks that need to be thoughtfully, carefully, and systematically considered.

6 CONCLUSION

We contribute an in-depth exploration of VFAIs to support aging in place, and implications for design and areas for future work for VFAIs tailored to older adults. We share the stories of five older adults who became Echo Show users over the duration of our study, and engaged in various aspects of the design process with us to imagine potential futures for these devices to serve older adults. Our design probes opened up possibilities for speculation that were difficult to imagine without them. Participating in our study and engaging with our design probes empowered participants to contribute ideas for new interface designs or improvements. We discuss design strategies for personalizing interactions to users' needs in real-time. Our work extends the existing literature surrounding the design of VFAIs to support aging in place, and calls for more research entailing their use for enabling continuity of care in older adults' homes.

ACKNOWLEDGMENTS

This research was funded by NSF Awards #2026577 and #1700832. AC was additionally supported by a Digital Life Initiative Doctoral Fellowship, and a PRISM Baker Postdoctoral Fellowship. We sincerely thank all our study participants and our collaborators. We gratefully acknowledge our collaborators who contributed to the data analysis (Jessie Taft), technical explorations of the health design probe (Anurima Sharma), device deployment (Hyein Baek) and the senior center directors who helped us connect with our participants. In addition, we thank Dr. John Brownstein, and Sarah Lindenauer (Boston Children's Hospital); Monica Allison, Dr. Fernanda Polubriaginof, and Dr. Dan Stein (Memorial Sloan Kettering Cancer Center); Dr. Elaine Wethington, Dr. Cary Reid, and Dr. Sara Czaja (Weill Cornell); and Dr. Jose Luchsinger, Dr. Jennifer Manly, and Dr. Adriana Arcia (Columbia) for their feedback on the concept of the design probe for health data reporting, connections, and guidance. We are especially grateful to Martin Olson at StoryCorps for his advice and mentorship in the development of the design probe for positive reminiscing. Finally, we would like to thank Dr. Kristen Fessele for her support with this work, and helping us move it forward in the medical world!

REFERENCES

[1] [n.d.]. Aiva Health gets investment from Google, Amazon - Bizwomen. Retrieved 2021-09-02 from https://www.bizjournals.com/bizwomen/news/latest-news/2018/10/healthcare-startup-aiva-health-gets-investments.html?page=all.

[2] [n.d.]. Endel raises $5M to create personalized 'sound environments' that improve productivity and sleep | TechCrunch. Retrieved 2021-09-02 from https://techcrunch.com/2020/09/23/endel-series-a/.

[3] [n.d.]. Great Questions – StoryCorps. Retrieved 2021-09-07 from https://storycorps.org/participate/great-questions/.

[4] [n.d.]. MyndYou - Health Predictive Analytics - MyndYou. Retrieved 2023-02-01 from https://myndyou.com/.

[5] [n.d.]. Virtual Assistant Technology - Statistics & Facts | Statista. Retrieved 2021-09-02 from https://www.statista.com/topics/5572/virtual-assistants/.

[6] Sarah Abdi, Luc de Witte, and Mark Hawley. 2020. Emerging technologies with potential care and support applications for older people: review of gray literature. *JMIR aging* 3, 2 (2020), e17286.

[7] Anneliese Arnold, Stephanie Kolody, Aidan Comeau, and Antonio Miguel Cruz. 2022. What does the literature say about the use of personal voice assistants in older adults? A scoping review. *Disability and Rehabilitation: Assistive Technology* (2022), 1–12.

[8] Shari S Bassuk, Thomas A Glass, and Lisa F Berkman. 1999. Social disengagement and incident cognitive decline in community-dwelling elderly persons. *Annals of internal medicine* 131, 3 (1999), 165–173.

[9] Beverly Beisgen and Marilyn Kraitchman. 2003. *Senior centers: Opportunities for successful aging*. Springer Publishing Company.

[10] Erin Beneteau, Olivia K Richards, Mingrui Zhang, Julie A Kientz, Jason Yip, and Alexis Hiniker. 2019. Communication breakdowns between families and Alexa. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[11] Frank Bentley, Chris Luvogt, Max Silverman, Rushani Wirasinghe, Brooke White, and Danielle Lottridge. 2018. Understanding the long-term use of smart speaker assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–24.

[12] Eva Brandt, Thomas Binder, Lone Malmborg, and Tomas Sokoler. 2010. Communities of everyday practice and situated elderliness as an approach to co-design for senior interaction. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. 400–403.

[13] Virginia Braun and Victoria Clarke. 2021. *Thematic analysis: A practical guide*. Sage.

[14] Pierce C. Upadhyay P. Park L. Brewer, R.N. 2021. A Question of Access: Exploring the Perceived Benefits and Barriers of Intelligent Voice Assistants for Improving Access to Consumer Health Resources Among Low-Income Older Adults. *ACM Transactions on Interactive Intelligent Systems (TIIS)* 1 (2021), 39.

[15] Robin Brewer. 2016. Connecting older adults through voice-based interfaces. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. 131–134.

[16] Robin Brewer, Casey Pierce, Pooja Upadhyay, and Leeseul Park. 2022. An empirical study of older adult's voice assistant use for health information seeking. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 12, 2 (2022), 1–32.

[17] Fred B Bryant, Colette M Smart, and Scott P King. 2005. Using the past to enhance the present: Boosting happiness through positive reminiscence. *Journal of Happiness Studies* 6, 3 (2005), 227–260.

[18] US Census Bureau. [n.d.]. Older People Projected to Outnumber Children. Retrieved 2021-09-01 from https://www.census.gov/newsroom/press-releases/2018/cb18-41-population-projections.html.

[19] Philippe Cappeliez, Marilyn Guindon, and Annie Robitaille. 2008. Functions of reminiscence and emotional regulation among older adults. *Journal of Aging Studies* 22, 3 (2008), 266–272.

[20] Chen Chen, Janet G Johnson, Kemeberly Charles, Alice Lee, Ella T Lifset, Michael Hogarth, Alison A Moore, Emilia Farcas, and Nadir Weibel. 2021. Understanding Barriers and Design Opportunities to Improve Healthcare and QOL for Older Adults through Voice Assistants. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–16.

[21] Kyungjin Chung, Young Hoon Oh, and Da Young Ju. 2019. Elderly Users' Interaction with Conversational Agent. In *Proceedings of the 7th International Conference on Human-Agent Interaction*. 277–279.

[22] Leigh Clark, Nadia Pantidi, Orla Cooney, Philip Doyle, Diego Garaialde, Justin Edwards, Brendan Spillane, Emer Gilmartin, Christine Murad, Cosmin Munteanu, Vincent Wade, and Benjamin R. Cowan. 2019. What makes a good conversation? Challenges in designing truly conversational agents. *Conference on Human Factors in Computing Systems - Proceedings* (2019), 1–12. https://doi.org/10.1145/3290605.3300705 arXiv:1901.06525

[23] Cynthia F Corbett, Pamela J Wright, Kate Jones, and Michael Parmer. 2021. Voice-Activated Virtual Home Assistant Use and Social Isolation and Loneliness Among Older Adults: Mini Review. *Frontiers in Public Health* 9 (2021).

[24] Erin York Cornwell and Linda J Waite. 2009. Social disconnectedness, perceived isolation, and health among older adults. *Journal of health and social behavior* 50, 1 (2009), 31–48.

[25] Andrea Cuadra, Hyein Baek, Deborah Estrin, Malte Jung, and Nicola Dell. 2022. On Inclusion: Video Analysis of Older Adult Interactions with a Multi-Modal Voice Assistant in a Public Setting. In *International Conference on Information & Communication Technologies and Development (ICTD)*.

[26] Andrea Cuadra, Oluseye Bankole, and Michael Sobolev. 2021. Planning Habit: Daily Planning Prompts with Alexa. In *International Conference on Persuasive Technology*. Springer, 73–87.

[27] Andrea Cuadra, Yen-Hao Chen, Kae-Jer Cho, Deborah Estrin, and Armin Shahrokni. 2022. Introducing the v-RFA, a voice assistant-based geriatric assessment. *Journal of Geriatric Oncology* 13, 8 (2022), 1253–1255.

[28] Andrea Cuadra, Amy L Tin, Gordon Taylor Moffat, Koshy Alexander, Robert J Downey, Beatriz Korc-Grodzicki, Andrew J Vickers, and Armin Shahrokni. 2022. The association between perioperative frailty and ability to complete a web-based geriatric assessment among older adults with cancer. *European Journal of Surgical Oncology* (2022).

[29] Sara J Czaja, Walter R Boot, Neil Charness, and Wendy A Rogers. 2019. *Designing for older adults: Principles and creative human factors approaches*. CRC press.

[30] Sara J Czaja and Chin Chin Lee. 2002. Designing computer systems for older adults. In *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*. 413–427.

[31] Sara J Czaja and Chin Chin Lee. 2007. The impact of aging on access to technology. *Universal access in the information society* 5, 4 (2007), 341–349.

[32] Smit Desai and Jessie Chin. 2023. OK Google, Let's Learn: Using Voice User Interfaces for Informal Self-Regulated Learning of Health Topics among Younger and Older Adults. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–21.

[33] Anna Dickinson and Peter Gregor. 2006. Computer use has no demonstrated impact on the well-being of older adults. *International Journal of Human-Computer Studies* 64, 8 (2006), 744–753.

[34] Alan Dix. 2010. Human–computer interaction: A stable discipline, a nascent science, and the growth of the long tail. *Interacting with computers* 22, 1 (2010), 13–27.

[35] Andrew Ennis, Joseph Rafferty, Jonathan Synnott, Ian Cleland, Chris Nugent, Andrea Selby, Sharon McIlroy, Ambre Berthelot, and Giovanni Masci. 2017. A smart cabinet and voice assistant to support independence in older adults. In *International Conference on Ubiquitous Computing and Ambient Intelligence*. Springer, 466–472.

[36] Nancy M Gell, Dori E Rosenberg, George Demiris, Andrea Z LaCroix, and Kushang V Patel. 2015. Patterns of technology use among older adults with and without disabilities. *The Gerontologist* 55, 3 (2015), 412–421.

[37] James M Gill, Arch G Mainous III, and Musa Nsereko. 2000. The effect of continuity of care on emergency department use. *Archives of family medicine* 9, 4 (2000), 333.

[38] Kylie L Goodman and Christopher B Mayhorn. 2023. It's not what you say but how you say it: Examining the influence of perceived voice assistant gender and pitch on trust and reliance. *Applied Ergonomics* 106 (2023), 103864.

[39] Bruce Guthrie, John W Saultz, George K Freeman, and Jeannie L Haggerty. 2008. Continuity of care matters. *Bmj* 337 (2008).

[40] Jeannie L Haggerty, Robert J Reid, George K Freeman, Barbara H Starfield, Carol E Adair, and Rachael McKendry. 2003. Continuity of care: a multidisciplinary review. *Bmj* 327, 7425 (2003), 1219–1221.

[41] Christina Harrington, Aqueasha Martin-Hammond, and Kirsten E Bray. 2022. Examining Identity as a Variable of Health Technology Research for Older Adults: A Systematic Review. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–24.

[42] Christina N Harrington and Lisa Egede. 2023. Trust, Comfort and Relatability: Understanding Black Older Adults' Perceptions of Chatbot Design for Health Information Seeking. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[43] Christina N Harrington, Radhika Garg, Amanda Woodward, and Dimitri Williams. 2022. "It's Kind of Like Code-Switching": Black Older Adults' Experiences with a Voice Assistant for Health Information Seeking. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.

[44] Hsiu-Fang Hsieh and Jing-Jy Wang. 2003. Effect of reminiscence therapy on depression in older adults: a systematic review. *International journal of nursing studies* 40, 4 (2003), 335–345.

[45] Esther Iecovich. 2014. Aging in place: From theory to practice. *Anthropological notebooks* 20, 1 (2014).

[46] Keith Johnstone. 2012. *Impro: Improvisation and the theatre*. Routledge.

[47] Pegah Karimi, Kallista Ballard, Pooja Vazirani, Ravi Teja Narasimha Jorigay, and Aqueasha Martin-Hammond. 2022. Designing Conversational Assistants to Support Older Adults' Personal Health Record Access. In *Pervasive Computing Technologies for Healthcare: 15th EAI International Conference, Pervasive Health 2021, Virtual Event, December 6-8, 2021, Proceedings*. Springer, 253–271.

[48] Yukta Karkera, Barsa Tandukar, Sowmya Chandra, and Aqueasha Martin-Hammond. 2023. Building Community Capacity: Exploring Voice Assistants to Support Older Adults in an Independent Living Community. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[49] Ngaire Kerse, Stephen Buetow, Arch G Mainous, Gregory Young, Gregor Coster, and Bruce Arroll. 2004. Physician-patient relationship and medication compliance: a primary care investigation. *The Annals of Family Medicine* 2, 5 (2004), 455–461.

[50] Shahedul Huq Khandkar. 2009. Open coding. *University of Calgary* 23 (2009), 2009.

[51] Sunyoung Kim and Abhishek Choudhury. 2021. Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study. *Computers in Human Behavior* (2021), 106914.

[52] Robert Kraut and Moira Burke. 2015. Internet use and psychological well-being: Effects of activity and audience. *Commun. ACM* 58, 12 (2015), 94–100.

[53] Amanda Lazar, Alisha Pradhan, Ben Jelen, Katie A. Siek, and Alex Leitch. 2021. Studying the Formation of an Older Adult-Led Makerspace. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.

[54] Chaiwoo Lee, Maryam FakhrHosseini, Julie Miller, Taylor R Patskanick, and Joseph F Coughlin. 2019. The Oldest Olds' Perceptions of Social Robots. In *International Conference on Human-Computer Interaction*. Springer, 405–415.

[55] Stephen Lindsay, Daniel Jackson, Guy Schofield, and Patrick Olivier. 2012. Engaging older people using participatory design. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1199–1208.

[56] Na Liu, Jiamin Yin, Sharon Swee-Lin Tan, Kee Yuan Ngiam, and Hock Hai Teo. 2021. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. *Journal of the American Medical Informatics Association* (2021).

[57] Gale M Lucas, Albert Rizzo, Jonathan Gratch, Stefan Scherer, Giota Stratou, Jill Boberg, and Louis-Philippe Morency. 2017. Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. *Frontiers in Robotics and AI* 4 (2017), 51.

[58] Arch G Mainous, Richard Baker, Margaret M Love, Denis P Gray, and James M Gill. 2001. Continuity of care and trust in one's physician: evidence from primary care in the United States and the United Kingdom. *Fam Med* 33, 1 (2001), 22–27.

[59] Aqueasha Martin-Hammond, Sravani Vemireddy, and Kartik Rao. 2019. Exploring older adults' beliefs about the use of intelligent assistants for consumer health information management: A participatory design study. *JMIR aging* 2, 2 (2019), e15381.

[60] Tracy L Mitzner, Julie B Boron, Cara Bailey Fausset, Anne E Adams, Neil Charness, Sara J Czaja, Katinka Dijkstra, Arthur D Fisk, Wendy A Rogers, and Joseph Sharit. 2010. Older adults talk technology: Technology usage and attitudes. *Computers in human behavior* 26, 6 (2010), 1710–1721.

[61] Phani Nallam, Siddhant Bhandari, Jamie Sanders, and Aqueasha Martin-Hammond. 2020. A Question of Access: Exploring the Perceived Benefits and Barriers of Intelligent Voice Assistants for Improving Access to Consumer Health Resources Among Low-Income Older Adults. *Gerontology and Geriatric Medicine* 6 (2020), 2333721420985975.

[62] Katherine O'Brien, Anna Liggett, Vanessa Ramirez-Zohfeld, Priya Sunkara, and Lee A Lindquist. 2020. Voice-Controlled Intelligent Personal Assistants to Support Aging in Place. *Journal of the American Geriatrics Society* 68, 1 (2020), 176–179.

[63] Ihudiya Finda Ogbonnaya-Ogburu, Angela DR Smith, Alexandra To, and Kentaro Toyama. 2020. Critical Race Theory for HCI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–16.

[64] Young Hoon Oh, Kyungjin Chung, Da Young Ju, et al. 2020. Differences in Interactions with a Conversational Agent. *International Journal of Environmental Research and Public Health* 17, 9 (2020), 3189.

[65] Norm O'Rourke, Philippe Cappeliez, and Amy Claxton. 2011. Functions of reminiscence and the psychological well-being of young-old and older adults over time. *Aging & Mental Health* 15, 2 (2011), 272–281.

[66] Alisha Pradhan, Leah Findlater, and Amanda Lazar. 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 (2019), 1–21.

[67] Alisha Pradhan and Amanda Lazar. 2020. Voice Technologies to Support Aging in Place: Opportunities and Challenges. *Innovation in Aging* 4, Suppl 1 (2020), 317.

[68] Alisha Pradhan, Amanda Lazar, and Leah Findlater. 2020. Use of intelligent voice assistants by older adults with low technology use. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 4 (2020), 1–27.

[69] Valeria Righi, Sergio Sayago, and Josep Blat. 2017. When we talk about older people in HCI, who are we talking about? Towards a 'turn to community'in the design of technologies for a growing ageing population. *International Journal of Human-Computer Studies* 108 (2017), 15–31.

[70] Lazlo Ring, Barbara Barry, Kathleen Totzke, and Timothy Bickmore. 2013. Addressing loneliness and isolation in older adults: Proactive affective agents provide better support. In *2013 Humaine Association conference on affective computing and intelligent interaction*. IEEE, 61–66.

[71] Yvonne Rogers and Gary Marsden. 2013. Does he take sugar? Moving beyond the rhetoric of compassion. *interactions* 20, 4 (2013), 48–57.

[72] Sergio Sayago, Barbara Barbosa Neves, and Benjamin R Cowan. 2019. Voice assistants and older people: some open issues. In *Proceedings of the 1st International Conference on Conversational User Interfaces*. 1–3.

[73] Hanna Schneider, Malin Eiband, Daniel Ullrich, and Andreas Butz. 2018. Empowerment in HCI-A survey and framework. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[74] Armin Shahrokni, Amy Tin, Robert J Downey, Vivian Strong, Sanam Mahmoudzadeh, Manpreet K Boparai, Sincere McMillan, Andrew Vickers, and Beatriz Korc-Grodzicki. 2017. Electronic rapid fitness assessment: a novel tool for preoperative evaluation of the geriatric oncology patient. *Journal of the National Comprehensive Cancer Network* 15, 2 (2017), 172–179.

[75] Jaisie Sin, Heloisa Candello, Leigh Clark, Benjamin R Cowan, Minha Lee, Cosmin Munteanu, Martin Porcheron, Sarah Theres Völkel, Stacy Branham, Robin N Brewer, et al. 2023. CUI@ CHI: Inclusive Design of CUIs Across Modalities and Mobilities. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–5.

[76] Jaisie Sin, Dongqing Chen, Jalena G Threatt, Anna Gorham, and Cosmin Munteanu. 2022. Does Alexa Live Up to the Hype? Contrasting Expectations from Mass Media Narratives and Older Adults' Hands-on Experiences of Voice Interfaces. In *Proceedings of the 4th Conference on Conversational User Interfaces*. 1–9.

[77] Jaisie Sin, Cosmin Munteanu, Dongqing Chen, and Jalena Threatt. 2022. Avoiding mixed messages: research-based fact-checking the media portrayals of voice user interfaces for older adults. *Human–Computer Interaction* (2022), 1–24.

[78] Brodrick Stigall, Jenny Waycott, Steven Baker, and Kelly Caine. 2019. Older adults' perception and use of voice user interfaces: A preliminary review of the computing literature. *ACM International Conference Proceeding Series* (2019), 423–427. https://doi.org/10.1145/3369457.3369506

[79] Arthur A Stone and Saul Shiffman. 1994. Ecological momentary assessment (EMA) in behavorial medicine. *Annals of behavioral medicine* (1994).

[80] Milka Trajkova and Aqueasha Martin-Hammond. 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*. 1–13.

[81] Pooja Upadhyay, Sharon Heung, Shiri Azenkot, and Robin N Brewer. 2023. Studying Exploration & Long-Term Use of Voice Assistants by Older Adults. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–11.

[82] Carl Van Walraven, Natalie Oake, Alison Jennings, and Alan J Forster. 2010. The association between continuity of care and outcomes: a systematic and critical review. *Journal of evaluation in clinical practice* 16, 5 (2010), 947–956.

[83] Laura Pfeifer Vardoulakis, Lazlo Ring, Barbara Barry, Candace L Sidner, and Timothy Bickmore. 2012. Designing relational agents as long term social companions for older adults. In *International conference on intelligent virtual agents*. Springer, 289–302.

[84] John Vines, Gary Pritchard, Peter Wright, Patrick Olivier, and Katie Brittain. 2015. An age-old problem: Examining the discourses of ageing in HCI and strategies for future research. *ACM Transactions on Computer-Human Interaction (TOCHI)*

22, 1 (2015), 1–27.

[85] Alexandra Voit, Jasmin Niess, Caroline Eckerth, Maike Ernst, Henrike Weingärtner, and Paweł W Woźniak. 2020. 'It's not a romantic relationship': Stories of Adoption and Abandonment of Smart Speakers at Home. In *19th International Conference on Mobile and Ubiquitous Multimedia*. 71–82.

[86] Kerryellen G Vroman, Sajay Arthanat, and Catherine Lysack. 2015. "Who over 65 is online?" Older adults' dispositions toward information communication technology. *Computers in Human Behavior* 43 (2015), 156–166.

[87] Nicole Wagner, Khaled Hassanein, and Milena Head. 2010. Computer use by older adults: A multi-disciplinary review. *Computers in human behavior* 26, 5 (2010), 870–882.

[88] Zhuxiaona Wei and James A Landay. 2018. Evaluating speech-based smart devices using new usability heuristics. *IEEE Pervasive Computing* 17, 2 (2018), 84–96.

[89] Julika Welge and Marc Hassenzahl. 2016. Better than human: about the psychological superpowers of robots. In *International Conference on Social Robotics*. Springer, 993–1002.

[90] Elaine Wethington, Christopher Eccleston, Geri Gay, Rachael Gooberman-Hill, Patricia Schofield, Elizabeth Bacon, Wen Dombrowski, Robert Jamison, Max Rothman, Lauren Meador, et al. 2018. Establishing a research agenda on mobile health technologies and later-life pain using an evidence-based consensus workshop approach. *The Journal of Pain* 19, 12 (2018), 1416–1423.

[91] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 493–502.

[92] Tamara Zubatiy, Kayci L Vickers, Niharika Mathur, and Elizabeth D Mynatt. 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*. 1–15.