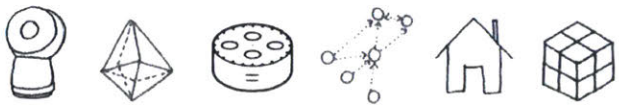
Talking Machines

Democratizing the design of voice-based agents for the home

NIKHITA SINGH



B.S. in Industrial Engineering & Operations Research
University of California-Berkeley 2014

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology.

JUNE 2018

Signature of Author:
Signature redacted
Program in Media Arts & Sciences,
May 15, 2018

Certified by:
Signature redacted
Cynthia Breazeal, Associate
Professor of Media Arts & Sciences,
Thesis Supervisor

Accepted by
Signature redacted
Tod Machover, Professor of Media
Arts & Sciences
Academic Head, Program in Media
Arts and Sciences

Talking Machines

Democratizing the design of voice-based agents for the home

NIKHITA SINGH

B.S. in Industrial Engineering & Operations Research
UC Berkeley 2014

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology.

ABSTRACT

Embodied voice-based agents, such as Amazon Echo, Google Home, and Jibo, are becoming increasingly present in the home environment. For most people, these agents represent their first experience living with artificial intelligence in such private and personal spaces. However, little is known about people's desires, preferences, and boundaries for these technologies. This thesis shares insights, learnings, methods, and tools from a journey with 69 children, adults, and older adults to help democratize the design of voice-based agents for the home.

In the first study, participants *interact with* and *discover* various voice-based agents to capture first impressions of the technology. In the second study, participants engage in long-term encounters with agents in their home to *experience* and *reflect* upon their preferences, desires and boundaries for these devices. Qualitative and quantitative data from interview transcripts, card sorting, and deployed cultural and technology probes is used to identify agent action preferences, socio-technical themes, daily usage trends, personality preferences, and future "wishes" for agents. This work culminates with participants *designing* their dream agents for the home through a structured ideation process. Throughout this work, a series of participatory design tools and methods are developed, iterated upon, and implemented to create a language of engagement with participants. These methods and tools are shared as an *open-source design kit* for others seeking to explore the domain.

Cynthia Breazeal, Thesis Supervisor
Associate Professor of Media Arts & Sciences

Talking Machines

Democratizing the design of voice-based agents for the home

NIKHITA SINGH

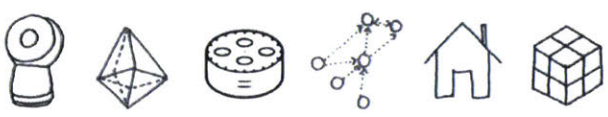
B.S. in Industrial Engineering & Operations Research
UC Berkeley 2014

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology.

Thesis Supervisor:
Signature redacted
Cynthia Breazeal, Associate Professor of Media Arts & Sciences

Thesis Reader:
Signature redacted
Neri Oxman, Associate Professor of Media Arts & Sciences

Thesis Reader:
Signature redacted
Rosalind Picard, Professor of Media Arts & Sciences



To the two people who have always believed in me and what I do.

Dadi (Grandma): What can it do?

Me: Show you pictures, talk to you, remind you of things, call me, answer questions. That kind of stuff.

Dadi: What does it eat?

Me: Battery lol.

Dadi: So it's like a pet, but better and you don't have to feed it? Everyone should have one!

Acknowledgements

One of my old mentors once said, "*We're fascinated by robots because they are reflections of ourselves*"—Ken Goldberg. I came to the lab searching for cool robots. In the process, I was inspired by the humanity, passion, and crazy ideas of people helping to design our future.

To mentors:
Cynthia, the first time you interviewed me for the lab, we sat in your living room on your couch, drank tea, and chatted about robots. I left that meeting in awe of how someone with such an incredible vision could be so humble, approachable, and real. Thank you for letting me find my voice through many random explorations, distractions and crazy ideas. I am immensely grateful for your trust, guidance and encouragement over the last two years.

Roz, every time we meet, I am inspired by the way you think, your commitment to students, and your ability to ask critical questions. Thank you for the many wonderful discussions that have challenged and shaped my thinking. Neri, thank you for pushing me to think deeper and find comfort in ambiguity. Your ability to understand, synthesize, and push ideas is awe inspiring. I am grateful for your beautiful energy and wisdom.

To the Personal Robots Group:
To all the wonderful members of the robots family, thank you for pushing my thinking and sharing in my passions. Polly, I will miss your stories, hugs and love. Anastasia, thank you for being my partner in crime and both enabling and engaging with my love for all things pretty. Every time I see a laminator, tomato and cheese, or find my letters escaping me, I will remember our adventures with people and robots fondly.

Hae Won, my time at the lab would not have been the same

Funding Sources:
This work was supported by the ICT R&D program of MSIP/IITP. in addition to Bose funding.

without you. Every time I came to your office to see if you had a *"few minutes to chat"*, you made yourself available (despite the fact that few minutes usually became an hour). Thank you for the countless whiteboard brainstorm sessions, pushing me to persevere through obstacles, and being an incredible friend and mentor. I hope that I have the honor of working with you again in the future.

Pedro and Ishaan, there is nothing I can say here that fully captures my gratitude for what you both have given me in the last two years. Thank you for sticking by me, always lending an ear regardless of the time of day, and being my family away from home. I will always remember our beers at Muddy's, sunrise conversations, and late night work sessions. Here is to *many, many,* more adventures together regardless of where we end up.

Finally, Safinah and Randi, thank you for all the impromptu life discussions and both encouraging and joining in my crazy.

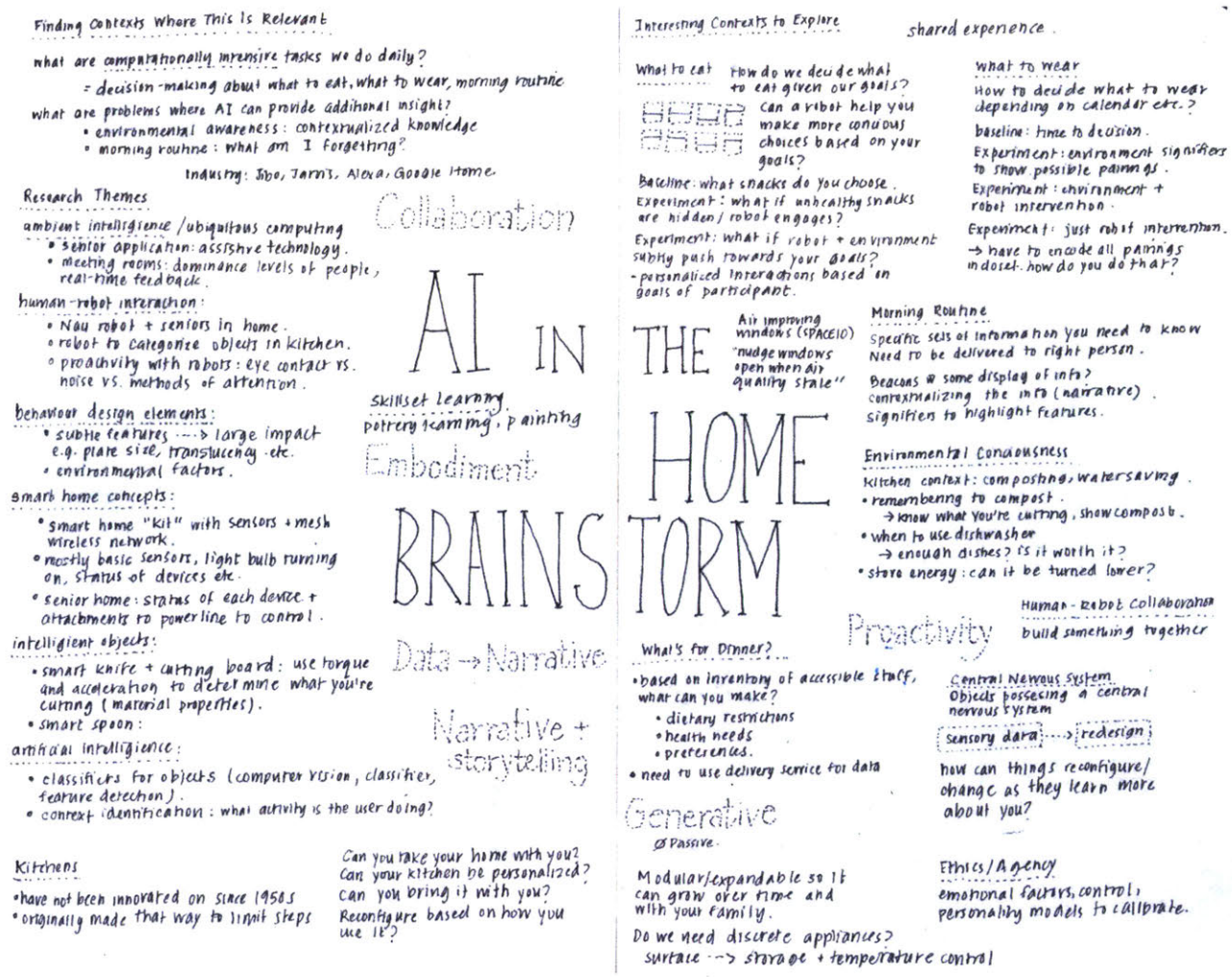
To friends and family:

Lynna, Shane, Rodrigo, Anmol, Kiron, Chrisoula and Anna, thank you for listening to my *many* rants and musings about robots and keeping me sane through these two years. There are very few friends that take long trains to see you for just a few hours, drive away to random towns with you to escape, call you everyday, and show up with cupcakes and wine whenever you need them. I don't know what I would do without each of you.

To my brother, Nish, thank you for always reminding me that there is more to life than robots and consistently inspiring me with your humor, candor, and authenticity. I am an *incredibly* lucky sister to have you around (and now you have this immortalized in print, uh oh).

Mumma, over the years I have come to deeply appreciate and admire your honesty, beautiful nature, and unwavering faith in the face of challenges. Thank you for always supporting my unusual life decisions, challenging my perspectives, and always striving to learn more about my mind and world.

Papa, it is difficult to capture all the ways in which you have inspired me in a few words. I still remember tinkering in the basement together, building science fair projects, and trying to copy your perfect handwriting. Your hard work, creativity, and sacrifice has always pushed me to tirelessly pursue my dreams. Thank you for all the *long* phone discussions, unwavering faith in my potential, and helping me find beauty in this world.



First project brainstorm page from notebook.

Table of Contents

Part I
Overview & Background

This thesis lives at the intersections of technology and design, people and agents, ambiguity and artificial intelligence. In Part I, (1) an overview of the central questions and themes surrounding the design of voice-based agents for the home is provided, (2) prior work in the domain is explored to contextualize the chosen approach and (3) the evolution of methods and tools utilized in this work is described. These chapters are outlined below.

PART I

1.0 Introduction

The idea of talking to machines has longstanding roots in science fiction—from HAL 9000 in *2001: A Space Odyssey* to Samantha in *Her*. Today, people are inviting *talking machines* into their homes at an unprecedented rate. In the United States alone, 43.5 million people currently own a smart-speaker in their home environment (Voicebot.AI Report, 2018).

The theoretical appeal of voice user interfaces is simple—more natural, intuitive, convenient, and efficient. In addition, the use of voice has the capability to offer an accessible interface for those who are visually, physically, or cognitively impaired. However, in practice, these voices in people's lives exist in the deep gulf between what they *want* them be and what they are *capable of* today. The experience of utilizing voice-based agents in current state involves adapting to clunky language and dealing with inconsistency and misunderstanding. Despite these challenges, people are open to and intrigued by the *human-agent relationship* and its potential to affect their lives.

Voice-based agents intended for the domestic environment such as Amazon Echo, Google Home, and Jibo embody a unique experience of technology *sharing personal space* with people. Unlike other technologies in the home that act as tools, these agents are designed to create bi-directional interaction through proactive behavior and personalization. The experiences enabled by these devices to accomplish this goal, however, vary significantly—some providing a more "assistant-like" *transactional* experience, with others drawing on more *relational* qualities through the integration of personalities and social expressivity.

Despite the proliferation of voice-based agents, the mechanics and impact of their interaction with people in the home context are yet to be understood. As people welcome

agents into their homes, these devices rub against the routines, rituals, and systems that compose the private, comfortable, and intimate nature of personal spaces. In addition, the *sharing of day-to-day life* with such agents elicits powerful questions surrounding the role, perceived relationship, and socio-emotional effects of the technology in people's lives.

This thesis presents insights and learnings from a journey with 69 children, adults, and older adults to interact with, discover (Part II), experience, reflect upon (Part III), and design (Part IV) voice-based agents for the home environment. The core questions central to this work are:

1. What are people's preferences, desires and boundaries for voice-based agent technology in the home?
2. How do these preferences, desires, and boundaries evolve as people live with voice-based agents?
3. How do people envision the design of future voice-based agents for their home?

In exploring these questions, this thesis contributes to the fields of *human-computer interaction* and *human-centered design* in several key ways:

1. It provides an *intergenerational perspective* from children, adults, and older adults on functionalities, capabilities, and future designs of voice-based agents for the home.
2. It reflects a *long-term engagement* of participants, allowing for rich insights surrounding in-home experiences and changes in perspective over time.
3. It seeks to better characterize experiences around *living with personified artificial intelligence* in the home through the comparison of two different agent platforms (smart speaker verses social robot).
4. It shares a series of *tools and methods* for designers and technologists to further explore and develop the space.

Through these contributions, this work aims to **help democratize the design of voice-based agents for the home environment** and provide methods, insights, and tools for designers and technologists in the domain.

PART I

2.0 Background

This work lives at the intersections of human-computer interaction and design research methodologies. Firstly, the development, value, and limitations of voice user interfaces is considered. Secondly, the mechanics of human-machine communication are explored including the role of embodiment, multimodality, and behavioral realism. Finally, learnings from design research methodologies as well as studies of technologies in the home environment are highlighted to contextualize the chosen approach.

2.1 Voice user interfaces

Voice user interfaces (VUIs) enable people to interact with smart devices in a natural, intuitive manner through spoken language. (McTear, Callejas, & Griol, 2016; Porcheron, Fischer, Reeves, & Sharples, 2018). In literature, this term is often utilized interchangeably with conversational agents, speed-based natural user interfaces (NUIs) or intelligent personal assistants. Commercially available examples of such interfaces include Alexa in Amazon Echo and Assistant in Google Home. It is important to note the focus of this work on physically embodied voice user interfaces for the home to distinguish from mobile-based interfaces such as Siri or Cortana.

The notion of "talking to machines" has longstanding roots in science fiction (HAL 9000 in *2001: A Space Odyssey* to Samantha in *Her)* and human-computer interaction (HCI) research. Early visionary works in HCI such as JCR Licklide's *Man-Machine Symbiosis* (1960) and Reeves' and Nass' *The Media Equation* (1996) explored fundamental ideas that have inspired the development and evaluation of human-machine dialogue

In *Man-Machine Symbiosis* (1960), Licklide wrote with regards to speech production and recognition, "...there is a continuing interest in the idea of talking with computing machines. In large part, the interest stems from realization that one can hardly take a military commander or a corporation president away from his work to teach him to type. If computing machines are ever to be used directly by top-level decision makers, it may be worthwhile to provide communication via the most natural means, even at considerable cost."

systems for a range of context-settings such as eldercare (Portet, Vacher, Golanski, Roux, & Meillon, 2013; Torta et al., 2014), the home environment (Porcheron et al., 2018), healthcare (DeVault et al., 2014), and public spaces (Kopp. Gesellensetter, Krämer, & Wachsmuth, 2014). While early implementations of such systems enabled only simple, specific interactions, recent developments in speech processing, artificial intelligence, and access to user data have enabled higher task complexity and design (McTear et al. 2016). Additionally, voice user interfaces have extended beyond the mobile phone and taken on a wide range of embodied forms.

2.2 Talking machines

DIALOGUE AS INTERACTION

The value of natural language as an interaction medium between people and machines is often debated. Brennan (1998) argues that speech-based interfaces become most useful in higher complexity tasks where direct manipulation interfaces may become overloaded and difficult to use. Users feel less burdened and are able to delegate a portion of the task goal to an "intelligent" agent. Other work has also highlighted the value of natural language in tasks where a user is asking a question, seeking advice, or when prior state or context is relevant (Frohlich, 1997; Qvarfordt, 2004). Conversely, it can also be argued that the need for precise syntax and commands, combined with the unpredictability of speech interfaces renders them less appropriate as means of human-machine communication (Shneiderman, 1982; Shneiderman, 1992; Brennan, 1998).

Despite apparent tradeoffs, the proliferation of consumer voice user interfaces such as Apple's Siri, Microsoft's Cortana, Amazon's Alexa, and Google's Assistant has reached an all-time high with the promise of hands-free, convenient task management and access to information (Pieraccini 2012; Moore, 2014). Recent work studying the state of such commercially available voice user interfaces, however, highlights the deep "gulf" between user expectation and experience of such devices (Luger & Sullen, 2016). Despite impressive advances in speech processing, state-of-the-art systems lack the reliability and accuracy needed to transition users fully away from more traditional interfaces such as touch screens and keyboards

(Moore, 2014). To combat these challenges, Moore (2014) argues the need to look beyond speech alone and consider the communicative behaviors that compose the spoken *interaction*—nonverbal cues, gestures, prosodic cues, and context, by seeking inspiration from social psychology, developmental robotics and cognitive science. In doing so, however, it is critical to first consider whether interactions with such devices *should* replicate human-human dynamics—a question that is of particular relevance for the home environment where comfort, intimacy, and social relationships are paramount.

THE WAY WE TALK TO MACHINES

For decades, the ease and naturalness of human-human communication has been a source of inspiration and study for human-machine interaction. In *The Media Equation* (1996), Reeves and Nass showed that people respond to computer-based agents as social actors, emulating the dynamics of human-human interactions. This work demonstrated key ideas related to human-machine communication including: (1) people respond socially and naturally to media irrespective of experience, education, age, technology proficiency or cultures, (2) people attribute personality traits to computers such as a sense of humor, expertise, aggressiveness, and gender, (3) people prefer computers that express similar personality traits as them in comparison with those who do not, (4) people prefer computers that evolve to become more like them over time in comparison with those who maintain consistency in their similarity and (5) people like simplicity in the form of predictability and ease of use (Reeves & Naas, 1996).

In *The Media Equation* by Reeves and Naas (1996), "studies demonstrate that people are "polite" to computers; that they treat computers with female voices differently than "male" ones; that large faces on a screen can invade our personal space; and that on-screen and real-life motion can provoke the same physical responses."

Later work by Shechtman and Horowitz (2003) challenged this notion, asserting that while people do exhibit social reactions towards such agents, there are also key differences in communication dynamics. When interacting with another human, people tended to engage more, utilize more relationship-focused statements, invest more effort, and respond more assertively to assertions by their partner as compared to machines (Shechtman & Horowitz, 2003). It is important to note that participants in this study were made aware that they were talking to a computer and the interaction itself contained only basic social cues communicated through text-alone. As such, the generalizability of these findings to interactions between people

and socially-expressive agents is limited. Other works have also highlighted people's desire for humanistic qualities such as politeness and humor in agents, while emphasizing that subtle differences do exist in user expectation of the agent's interaction behavior (Wilks, 2010).

EMBODIMENT, MULTIMODALITY & BEING HUMAN

Building on these findings of human-agent interaction, researchers have argued that the perceived social influence of an autonomous agent is tightly linked to its behavioral realism (Von der Pütten, Krämer, Gratch, & Kang, 2010). As such, much of human-computer and human-robot interaction research in the domain has focused on leveraging embodiment and multimodality in system design in an effort to encapsulate and represent human-like behaviors.

One such example of an agent is Rea, an embodied conversational agent capable of both multimodal input comprehension and output expression (Cassell et al., 1999). Through the design and study of Rea, Cassell et al. demonstrate the importance of communicative functions—such as farewells, salutations, turn-taking—and nonverbal behaviors for convincing human-agent interaction (Cassell et al, 1999; Cassell et al., 2000; Cassell et al., 2001).

Moving beyond screen-based representations to physical forms, human-like behaviors have also been incorporated into embodied social agents or social robots. An iconic example of a social robot is Kismet, an expressive anthropomorphic robot capable of interacting physically, affectively, and socially with people in order to enable readable and expressive feedback, turn-taking, and proactive regulation of the interaction (Breazeal, 2002). Social robots such as Kismet have been shown to evoke meaningful interactions with users, interact with people in a natural and interpersonal manner, and exhibit social presence (Breazeal 2002; Lee, Park & Song, 2004; Jung, Lee, Takanishi, & Kobayashi, 2008).

Other representations of agents in this domain—both physical and software—have taken on a multitude of forms including functional, anthropomorphic, caricature or abstract, humanoid, robot objects, zoomorphic, and wearable.

2.3 Designing relationships between people & agents

The growing body of work in designing for human-agent interaction has raised central questions around the role of the agent and the corresponding modalities, expression and form best suited for context-specific scenarios. The design of voice-based interfaces for the home environment has broadly drawn on two approaches: *transactional* or *relational*. The differences between these exist not in *what* the agent does, but *how* the agent interacts with people.

TRANSACTIONAL VUIs

Transactional voice-based agents draw from the "virtual/digital assistant" or "virtual/digital butler" paradigm, supporting voice access to information and services as a utility. These agents act as an assistant for daily tasks and aim to learn about the user enough to act proactively on their behalf (Payr, 2012; Luger & Sullen, 2016). Examples of voice-based digital assistants can be found in Apple's Siri, Microsoft's Cortana and Google's Assistant in the phone. The dominant paradigm when interacting with these voice-controlled assistants involves the user initiating the interaction to order to gain access to information or services. Over time, such VUIs identify patterns about users in the background, allowing for improved personalization and proactivity in providing assistance. As such, their main focus is to prioritize maximal utility over interpersonal relationships, and act as a tool for the user, leading to an experience that is more device-like and functional in nature.

RELATIONAL VUIs

In contrast, another class of voice-based agents aims to extend the role beyond that of an assistant and build longer-term socio-emotional relationships with users (Bickmore & Picard, 2005). Such agents, referred to as "relational agents" or "companions", focus on acting as a collaborative partner or teammate to help the user achieve a goal. This is in contrast with transactional agents where utility is paramount. To accomplish this, relational agents express social and relational behaviors such as empathy, social dialogue, humor, and continuity behaviors to establish a social bond with users. Relational agents have been utilized for a range of applications including education, persuasion, health and

wellness, eldercare, business, and psychotherapy (Bickmore & Picard; 2005; Bickmore, Schulman & Yin, 2010). Commercial applications of these agents—such as AIBO, Furby, Pleo, and Tamagotchis—are largely embodied in caricature or zoomorphic forms, pulling on people's desire to nurture and establish bonds with others.

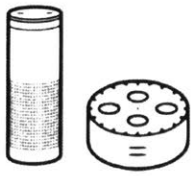
Similar to transactional agents, relational agents are capable of personalization and proactivity, but differ in their approach. Relational agents aim to learn about the user by emulating dynamics of human-human interactions. As such, the design of such agents often focuses on designing compelling personalities including likes and dislikes, opinions, and potentially even their own goals. By evolving these aspects of the technology, relational agents can take on the role of an assistant as well as a companion.

AGENTS FOR THE HOME

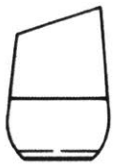
Currently available examples of such agents include Alexa in Amazon Echo, Google Assistant in Google Home and Jibo. This range of agents draws on various qualities of transactional and relational voice-user interfaces to create highly distinct experiences for users.

Amazon's Alexa and Google Home are both embodied as "smart speakers" for the domestic environment, with the goal of providing hands-free and convenient access to information, services, entertainment and messaging or calls. While both have physical forms, use voice as their primary mode of interaction, and provide visual feedback in the form of a light ring and dots respectively, their embodiment and social expression is largely functional and limited. Both devices offer significant utility through a range of skills focused on entertainment (e.g. jokes, music etc.), information access (e.g. weather, news), daily tasks (e.g. reminders, calendar) and social interaction (e.g. messaging or hands-free calling). However, much of the personalization and proactive behavior surrounding these skills must be enabled largely through user-driven input in the phone application.

Drawing on some elements of relational VUIs, both of these agents have personas including likes and dislikes, opinions and more. While Amazon Alexa and Google Home differ subtly in their expression of these personas, both have been designed as formal, adult females. Despite the existence of an inherent



Alexa in Amazon Echo was released in 2014 and has since been embodied in a range of forms including the Echo, Echo Dot, Echo Show, and Echo Spot. The Amazon Echo is intended as a hands-free assistant for the home and has the highest adoption of all commercially available embodied voice user interfaces for the home.



Google Assistant in Google Home was released in 2016 and has since been embodied in both the original and a Google Home Mini and Google Home Max form. It is intended as a hands-free assistant for the home and has the second largest adoption of all embodied voice user interfaces for the home.

Jibo was released in 2017 as a social robot for the home. The caricature robot is intended as a companion and has a range of expressive behaviors. Jibo can recognize faces and express social behaviors such as gaze, facial expressions, greetings, and humor. Jibo is currently in its early adoption stages and undergoing new updates.

persona, these agents do not focus on building relationships through socially expressive behavior (i.e. no greetings, farewells). Consequently, the experiences enabled by these agents are more transactional, functional, and similar to a personal assistant.

Smart speakers, such as those described above, emphasize the transactional utility offered by the agent while integrating relational aspects into the persona design. In contrast, Jibo, a voice-based agent embodied as a social robot, emphasizes being a *helpful companion.* While Jibo does provide utility—though less than Amazon's Alexa and Google Home—by supporting transactional queries (e.g. information, timers etc.), it *also* seeks to foster social bonds with people by creating interactions that are more overtly personal and personalized in nature. In order to accomplish this, Jibo leverages social-emotional behaviors such as gaze, facial expressions, greetings, humor, voice and face recognition and a friendly personality. Unlike the more simplistic personality of Amazon's Alexa and Google's Assistant, Jibo embodies a male personality that is informal, has humor, and designed to be a "part of the family". To support these qualities, Jibo also possesses additional modalities for (1) *user input* including visual perception and the ability to sense capacitive touch, as well as (2) *expression* including a screen for visual feedback and range of motors for body and head movements. The resulting experience is one that feels more alive and animate.

The range of agents described—Amazon Echo, Google Home, and Jibo—illustrate highly distinct approaches to designing voice-based user interfaces for the home. With voice-based agents entering homes at an unprecedented rate, it is more relevant than ever to consider the affordances of and desires for the technology in the home environment.

2.4 Technology and the home

Despite the growing adoption of voice-based agents in the present "smart home ecosystem", researchers are divided on the experiences they should create for users. One guiding framework argues that technology in the domestic environment should be calm, ubiquitous, proactive, and invisible (Weiser, 1991). Roger (2006) challenges this notion, arguing for a shift from *proactive technology* to *proactive people,* where technologies are designed not

to do things *for* people, but to engage them in a way that enables greater control of their lives and environment.

The private and intimate nature of the home poses methodological challenges for the design and study of HCI technologies. As Leppänen and Jokinen (2003) describe, the "home—as well as technology—evokes strong feelings". When technology seeks to penetrate the home, the sense of comfort, routines, traditions, and social norms that encapsulate the environment feel challenged, resulting in strong reservations and fear (Leppänen et al., 2003; De Graaf, Allouch & Van Dijk, 2016). Conversely, for HCI researchers, it is impossible to replicate the nuances of the home in the laboratory setting, thus making it critical to evaluate technologies in a deployed context.

Leppänen and Jokinen (2003) describe the conflict between the home and technology: "When talking about home, people see it ideally as a nest: home is safe, relaxing and comfortable. When talking about future technology, people's fears erupt. What if technology makes us cold and impersonal? Technology and home do not necessarily fit well together. That is why technology firms and smart home producers should think carefully about what kind of technological solutions they offer to consumers. How could technology be less like technology and more like home?"

Prior research evaluating agents for the home has relied on a range of methodologies including user interviews, focus groups, surveys, home visits, simulated lab studies, longer-term user studies, and short-term audio and video collection in the home (Forlizzi & DiSalvo, 2006; Klamer, Allouch & Heylen, 2010; Torta, Werner & Johnson, 2014; Cha, Forlizzi & Srinivasa, 2015; Luria et al., 2016; Luria, Hoffman & Zuckerman, 2017; De Graaf, Allouch & Van Dijk, 2017). While comparing a social robot to other smart-home interfaces including a voice-controlled speaker, touch-screen, and mobile application in a simulated lab environment, Luria et al. (2017) found that individuals perceived each interaction mode to have different affordances and disadvantages. The social robot was seen by participants as more engaging, enjoyable, and situationally aware but had lower perceived usability and sense of control. In contrast, while participants found the hands-free and ubiquitous control offered by the voice-controlled speaker compelling, they also expressed a lack of control and situational awareness using the interface.

Long-term studies of agents in the home have revealed that short-term acceptance of agents relies on engagement and easy of use, while long-term acceptance depends on functional-relevance to the user (De Graaf, Allouch & Van Dijk, 2017). Other studies—though focused specifically on the elderly population—have highlighted the need to (1) preserve a sense of autonomy, (2) consider security and privacy concerns, (3) design for long-term engagement rather than novelty and (4) consider the emotional responses people may exhibit towards agents

(Portet, Vacher, Golanski, Roux & Meillon, 2013; Torta,Werner & Johnson, 2014).

Broader awareness and adoption of voice-based agents is a new phenomenon. As such, comparison studies of interfaces and their long-term impacts in the home are limited. Existing works have also brought to light key challenges including (1) low sample sizes, (2) limited population diversity, (3) lack of deployable, robust technology systems, (4) user availability and adherence to protocol, (5) privacy needs of the home, and (6) limited generalizability of lab studies to the home context. In addition, researchers have highlighted user's reticence towards discussing social responses evoked by technology, calling for designers to take such reservations into account by building trust, practicing empathy, and involving design research practices in the research process (De Graaf et al., 2016). To combat the complex nature of the problem, researchers have argued for involving users earlier on in the design process to obtain deeper understanding of motivations and desires (Ogonowski, Ley, Hess, Wan & Wulf, 2013; De Graaf et al., 2016).

2.5 Drawing on design

Recent years have seen growing interest in designing technologies for the domestic environment, presenting core challenges related to the diversity of ages, desires, motivations, abilities, and limitations that must be integrated (Hutchinson et al. 2003). The new and complex goal of designing agents for the home environment calls for a shift in attitude from designing *for* users to designing *with* users. As such, methodologies and insights from participatory design provide a compelling basis for this exploration. "Participatory design" describes a set of "theories, practices, and studies" intended to involve users as full participants in the process of designing computer-based products and activities (Muller & Drain, 2003). This approach has been applied to several domains including software and hardware design, architecture, sustainability, graphic design, and medicine (Kensing & Blomberg, 1998; Hutchinson et al., 2003; Bødker, Kensing & Simonsen, 2004; Ogonowski et al., 2013).

When describing participatory design, Sanders (2002) asserts the importance of simultaneously exploring perspectives of "what people do, what what people say, and what people make" in

"Participatory experience is not simply a method or set of methodologies, it is a mindset and an attitude about people. It is the belief that all people have something to offer to the design process and that they can be both articulate and creative when given appropriate tools with which to express themselves" (Elizabeth Sanders, 2002).

order to understand their thoughts, feelings, and dreams towards a product, interface, or system. To facilitate such interactions, a range of design methods and tools have been utilized in research studies alongside interviews and discussion including those described below.

(1) *Cultural probes* are designed to elicit and record responses from people through the use of an artifact such as a map, diary, camera, or postcard (Gaver, Dunne & Pacenti, 1999). Cultural probes have been deployed in a range of studies to capture idea generation, inspiration, values, and desires in the design process (Blythe & Monk, 2002; Gaver et al. 2013; Wilson, Hargreaves & Hauxwell-Baldwin, 2015).

(2) *Technology* probes build on cultural probes, but focus on collecting data, testing, and eliciting reflections from users about a technology deployed in the real-world setting (Hutchinson et al., 2003). Such probes take the form of a deployed technological artifact, combined with analysis and reflection.

(3) *Games* draw on a "design-by-playing" approach and act as a means to better structure and improve communication between participants and facilitators for idea generation. Games have been used by design practitioners to create a language of engagement for stakeholders and reach vulnerable populations (Brandt & Messeter, 2004; Branco, Quental & Ribeiro, 2017).

(4) *Bags-of-Stuff* provides a low-fidelity prototyping approach where participants create physical models of technologies using a bag of simple materials such as paper, crayons, clay, string, and glue. This approach has been shown to be powerful for intergenerational design teams as a means to equalize differences in ability (Druin, 1999; Yip, Clegg et al. 2013).

(5) *Card Sorting* provides a tactile and engaging mechanism to explore how participants group items (concepts, terms, or features) into categories and relate them to each other (Hanington & Martin, 2010). Card sorting has been shown to be particularly useful when terminology associated with the technology in question may be vague or confusing to users (Seale, McCreadie, Turner-Smith & Tinker 2002; Hanington & Martin, 2010).

As shown, designing voice-based agents for the domestic environment is a complex and evolving domain, with much remaining to be studied and understood. The described participatory design methods provide inspiration for a means to democratize the design of these technologies by integrating stakeholders into the process.

PART I

3.0 Process & Tools

This thesis aims to democratize the design of voice-based agents for the home. The core questions central to this work are:

4. What are people's preferences, desires and boundaries for voice-based agent technology in the home?
5. How do these preferences, desires, and boundaries evolve as people live with voice-based agents?
6. How do people envision the design of future voice-based agents for their home?

In order to explore these ideas, a range of design tools were developed, iterated, and implemented alongside both quantitative and qualitative analysis. Participants—including children, adults, and older adults—were recruited from the local community and engaged as design partners throughout the process. The evolution of the process and tools is described in detail in this section.

6.1 Defining values & approach

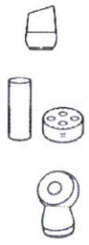
In order to conduct early explorations into the space, a range of techniques were utilized across populations including (1) shadow visits in people's homes, (2) visits and workshops with community organizations, and (3) an exploratory survey to gage people's current use of agents in the home environment. These explorations revealed that knowledge and acceptance of voice-based interfaces was often limited in the general population, and in particular among older adults. In fact, less than 20% of households currently owned a voice-based agent such as Amazon Alexa or Google Home based on an initially conducted survey of 150 households in North America.

When introducing voice-based agents for the home to people with no prior exposure, the technology was framed as:

"If there are so many of these devices around, then why have I never heard of it?"
– *Older adult*

"Alexa, what should we do about North Korea?"
– *Older adult's first interaction with Alexa*

"Google, could you do a better job than Donald Trump?"
– *Older adult's first interaction with Alexa*



"Voice-based agents provide a way for people to interact through voice (by speaking) to accomplish functional and social tasks. Think of it like an assistant or companion. It can help you find out information like the weather or news. Play music, Set reminders or timers. Turn on the lights or the thermostat. Or even provide a way to interact with others through messaging, photos, or calls."

"I'm not sure what I would do with it. I'm not disabled yet."
– *Adult*

"I want an Alexa for music. Google for googling, and Jibo for dancing. Then I'd also have a fidget spinner shaped one to come with me"
– *Child*

A collection of example videos and agent introductions were used to contextualize the space. Additionally, commercially available agents with voice-dialogue systems such as Amazon Alexa, Google Home, Siri, Cortana, Jibo, and others were used to provide an interactive experience. Insights from the initial exploration were used to develop a set of *principles* with which to approach the ensuing work. Defined here are five principles—some paralleling participatory design values—that guided the development and use of tools and methods in this work.

I. Technology design and development processes must be tangible and understandable to participants.

II. Participants must be treated as equal members of the design process, placing value and respect on their opinions and responses. This is aided by Principles III and IV.

III. A structured approach using design tools for engagement is necessary to provide a framework and encourage conversations around technology for populations who may lack knowledge about the technology.

IV. To fully engage in participatory design, participants must be equipped with a "language" of the technology. Given the recent emergence of voice-based agents, many people are unfamiliar with the technology. This provides them with a method of thoughtfully articulating their perspectives and increases their sense of value in participating.

V. It is critical to understand the technology with a holistic view of how it integrates into other relevant details of participants' lives including activities, relations to other people and community, and socio-emotional experiences.

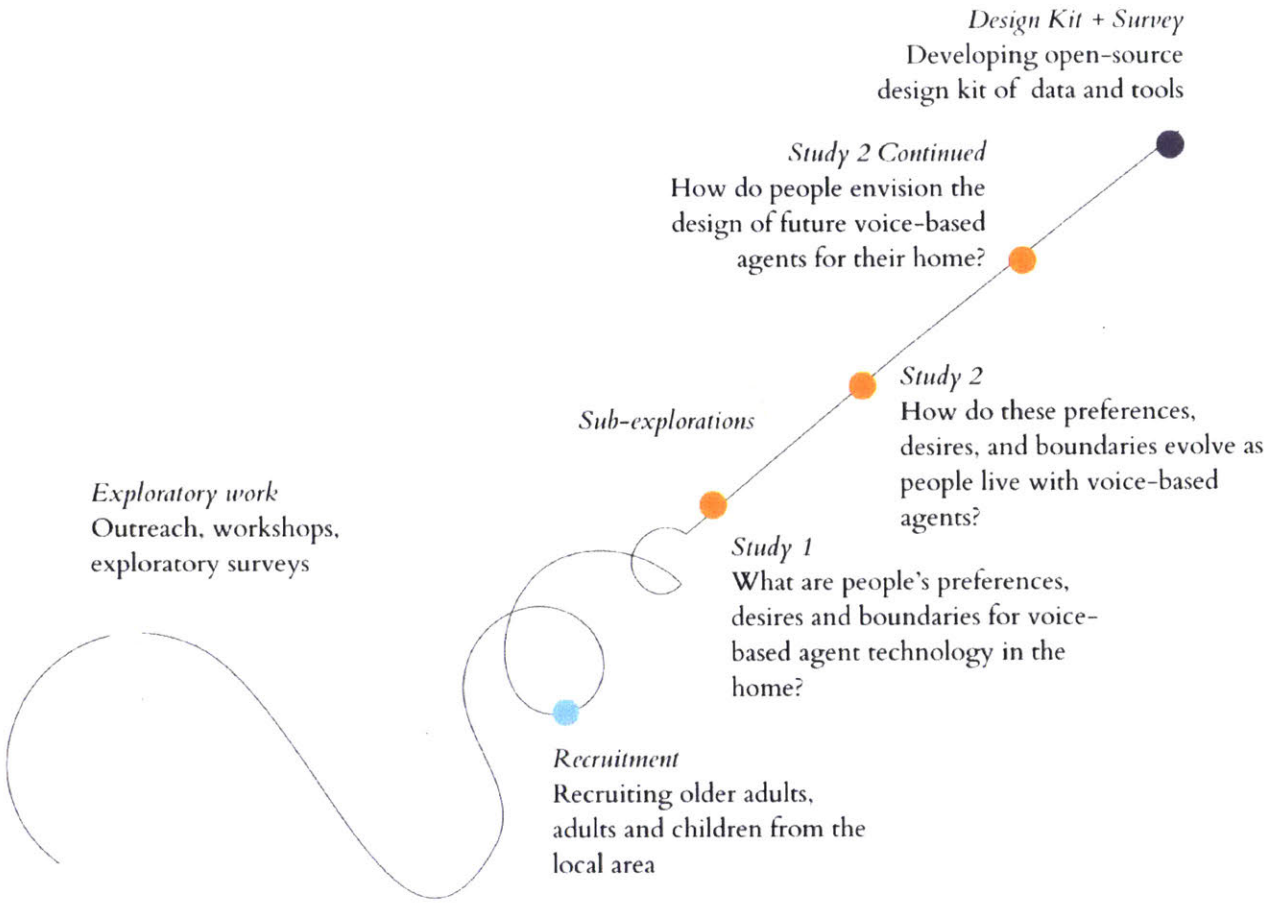


Figure 3.1: Overall trajectory and timeline of the thesis work.

6.2 Process and approach

This overarching trajectory of this work began with early explorations and recruitment of participants. Following this, two core studies were conducted in addition to smaller sub-explorations. The methods and tools utilized in this process converged in the creation of an open-source design kit for designing voice-based agents as well as a ~600-person survey to gather reflections from a broader community. This trajectory is visualized in Figure 3.1 above.

The objectives of the first study were to introduce participants to a diverse range of voice-based agents (*interact)* and capture their preferences, desires, and boundaries for the technology (*discover)*. This study utilized card sorting and semi-structured interview techniques. Following this study, a subset of the participants engaged in a longer-term study of voice-based

gents in their home. The objectives of this second study were to understand how participants' preferences, desires, and boundaries evolve after living with different voice-based agents in their home. Participants were given one of two highly distinct voice-based agents—Amazon Alexa or Jibo—as technology probes for a period of one month in their home *(experience)*. Those who currently owned an Amazon Alexa or Google Home were always given a Jibo to compare differences. During this deployment, cultural probes were embedded in participants' homes to capture in-context feedback and thoughts. Participants were brought back for a debrief session where they took part in a semi-structured interview and repeated the card sorting activity *(reflect)*. After these reflections, participants were guided through a structured ideation process to design their "dream voice-based agent" using the agent design cards *(design)*. These studies and the corresponding methodologies are shown as a design process in Figure 3.2. A total of 69 participants engaged in this work between the two studies including 19 children, 21 adults, and 29 older adults.

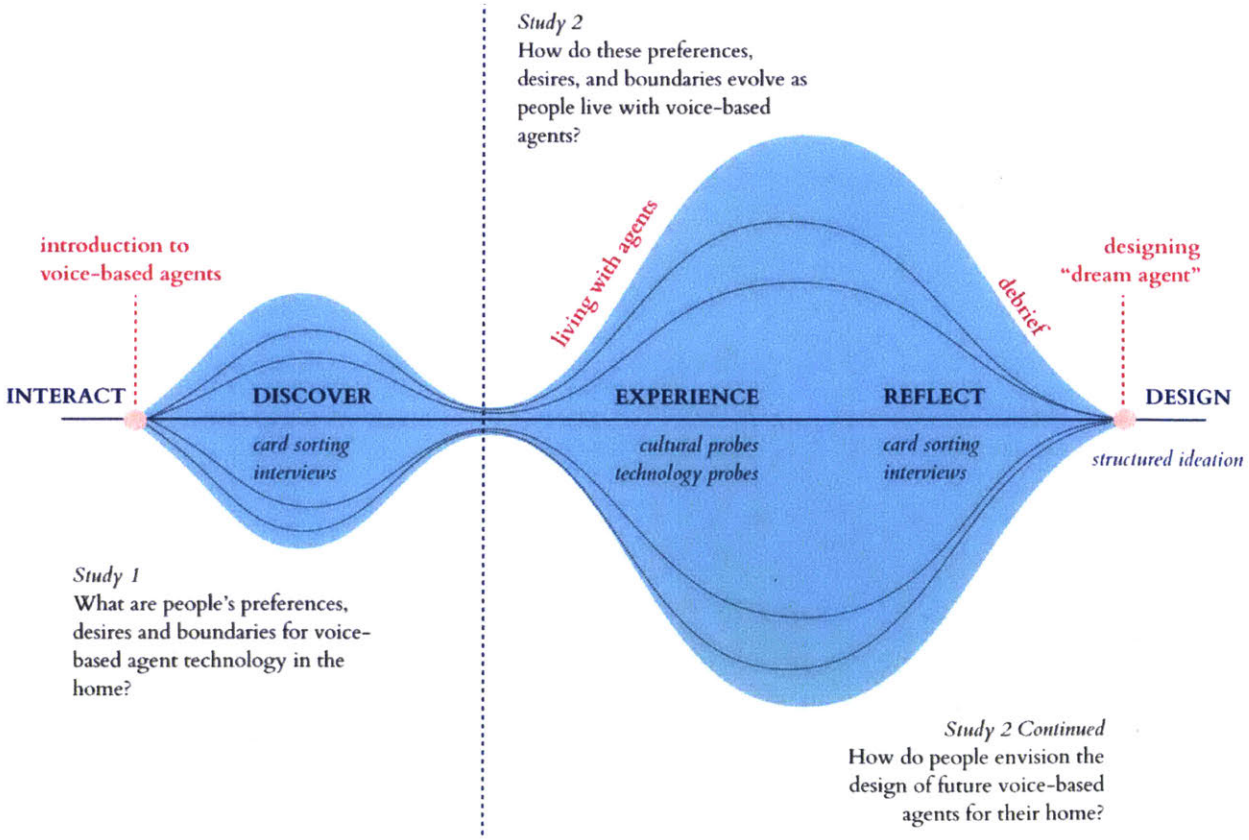
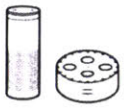
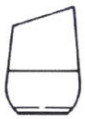


Figure 3.2: Diagram of overall methods including study 1 and 2. Corresponding questions and tools used are shown.

6.3 *Technology tools*

While Amazon Alexa, Google Home and Jibo are fully-deployed commercial products, it can be argued that given the new and evolving nature of these technologies, they can be utilized much like technology probes. Technology probes refer to installing a technology with a real world context, observing its use over time, and later reflecting on this use to capture users' preferences, desires and ideas for the technology (Hutchinson et al., 2003).

Amazon Alexa	
Description 	Amazon Alexa was released in 2014 and has since been embodied in a range of forms including the Echo, Echo Dot, Echo Show, and Echo Spot. The Amazon Echo is intended as a hands-free assistant for the home and has the highest adoption of all commercially available embodied voice user interfaces for the home.
Input modalities	Sound & speech
Output modalities	Sound & speech, visual (LED light ring)

Google Home	
Description 	Google Home was released in 2016 and has since been embodied in both the original, Mini, and Max form as a smart speaker. It is intended as a hands-free assistant for the home and has the second largest adoption of all embodied voice user interfaces for the home.
Input modalities	Sound & speech
Output modalities	Sound & speech, visual (LED dots)

Jibo	
Description 	*Jibo* was released in 2017 as a social robot for the home. The caricature robot embodies qualities of a relational agent and aims to be a helpful companion. The robot has a range of expressive behaviors to enable a more multimodal experience. Jibo can recognize faces and express social behaviors such as gaze, facial expressions, greetings, and humor.
Input modalities	Sound & speech, vision (camera), capacitive touch
Output modalities	Sound & speech, visual (LED light ring, screen), facial expressions, movement (body and head)

Table 3.1: Details about agents used as technology probes for studies.

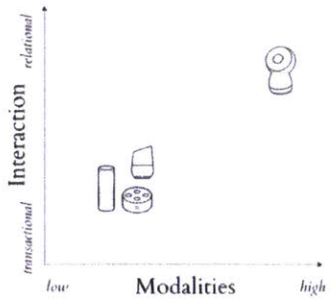


Figure 3.3: Amazon Alexa, Google Home, and Jibo placed relative to each other on dimensions of interaction type and number of modalities for interaction. It is important to note these separations are fluid and can evolve with time. This figure aims to capture the current state and show differences in intention and interaction with these agents.

Throughout studies with participants, these agents were positioned as exploratory tools rather than products for evaluation. Participants were explained that these voice-based systems for the home were simply examples and that the domain was constantly evolving with respect to agent forms, interactions, and experiences.

6.4 Design tools

CARD SORTING

Cards are realized as a way to support "focus shifts" and make it easier for participants to bring new perspectives and ideas to the design process (Halskov and Dalsgard, 2006). Card sorting as a technique has been shown to be particularly valuable in situations where the "language" around the technology is confusing or vague. As such, they provide a compelling and tactile means to engage the populations explored in this work. Additionally, the use of card sorting in this context allows the facilitators to guide users through the process in a fluid, yet structured way.

Leveraging these ideas, a set of 41 "agent action cards" were developed. Agent action cards were separated into six distinct categories, each exploring unique roles an agent could take on in a user's life. These are described in Table 3.2 below. Example actions represented by these cards are detailed in Figure 3.4.

Table 3.2: Description of agent action card categories.

Agent Action	Description of Category
Reminders	Cards in this category include tasks, events, or actions an agent could remind a user of. This category largely functional or assistant-type tasks for the agent. This reminder could be initiated by the user (i.e. "agent, do I have any meetings today?") or proactively by the agent.
Information	Cards in this category explored functional information about weather, news, etc. being provided by the agent. This category explored a largely functional or assistant-type tasks for the agent.
Suggestions	Cards in this category explored the idea of an agent providing a suggestion to the user to help build and maintain habits such as healthy eating, calling friends or family and more. This category aimed to explore users' perspectives on the agent having a goal system and taking on a more relational or companion role in their lives.
Share something	Cards in this category explored the idea of the agent taking on a more social role and having something to share with the user such as a playlist or joke.
Someone trying to reach you	Cards in this category introduced the idea of the agent being a means to for people to reach you for a phone call or update. A key differentiator of this category was the time-sensitive nature of the interaction.
Share something from someone else	Cards in this category explored the idea of the agent being a "social tool" to keep in touch with others. This included people using the agent as a means to share photos, videos etc. This category, along with the previous one, was meant to elicit perspectives on consolidation of technology, and the role of the agent as a facilitator of social connection.

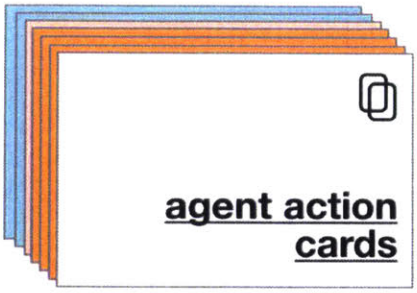
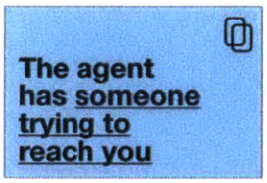


Figure 3.4: Card categories for card sorting activity used to engage participants in identifying their preferences and desires for agent actions. 41 agent actions were divided into six categories representing different roles the voice-based agent could play. Examples of cards within each category are described.

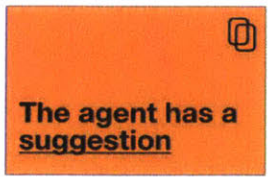
Examples of this could include reminders for birthdays, meetings, things to grab on your way out the door, grocery list, to-do-lists, medicine and more.

Examples of this could include the agent sharing music playlists you may like, jokes it thinks you may find funny, stories that you may enjoy, art or photography you may appreciate and more.

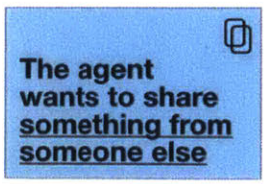
Examples of this could include daily weather, news, nutritional facts, where important objects are in the house, commute or traffic, TV programming, what a friend or family member is up to or when they are available for a call, and more.



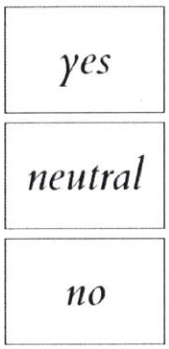
Examples of this could include someone calling you through the agent, someone sending you a status update (e.g. heading home), someone asking an important question through the agent (e.g. "are we out of tomatoes?") and more.



Examples of this could include a suggestion to drink water, meditate, learn something new, books to read, going for a walk, eating healthy, calling or meeting someone you haven't talked to in a while, social events in the area, or taking a nap.



Examples of this could include photos, video, songs or playlists, requests to play a game, stories or memories, messages from friends or family, a "hello" or "goodnight" to say they are thinking of you and more.



Actions in orange refer to functional goals that an agent can assist with.

Actions in [illegible] refer to socially-driven goals coming from the agent.

Actions in [illegible] refer to socially-driven goals initiated by another person through the agent.

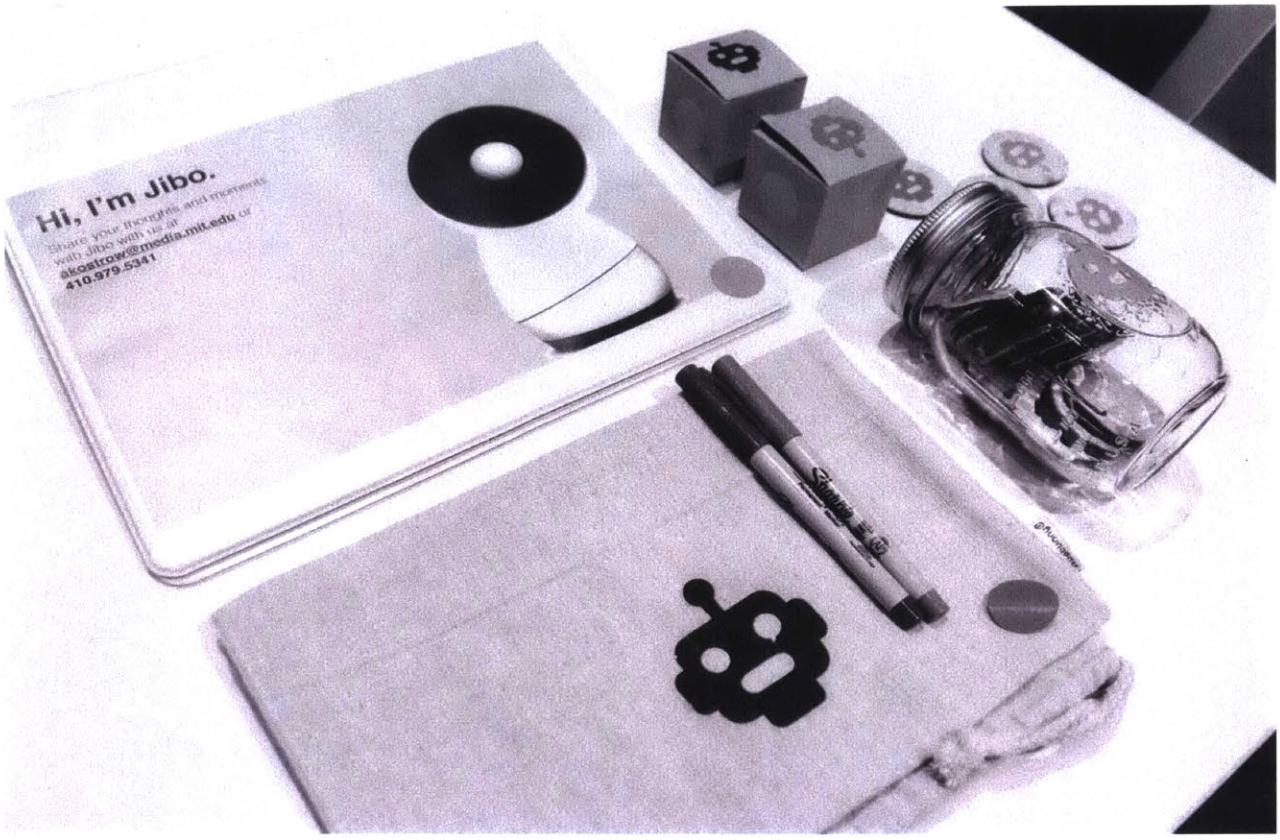


Figure 3.5: Photograph of cultural probe "kit" deployed with agents in the home to elicit and record responses from users.

CULTURAL PROBES

In order to elicit and capture responses from people during the agent's stay in their home, a range of cultural probes were designed and deployed including a (1) daily log book, (2) an "agent wish jar", and (3) mechanisms for free-form feedback.

The *daily log book* was designed to provide the family with prompts to guide their experience with the agent in the home. The log book included daily sheets with a range of functional social, and entertainment tasks the agents could do. Chosen tasks were consistent with those often assessed when studying voice-based technologies (López, Quesada & Guerrero, 2018). Simple words were used as prompts rather than detailed descriptions in order to encourage participants to self-discover the appropriate language to use with the agent and enable more natural interactions. Echoing the approach used with card sorting, each action was associated with a "enjoyed", "neutral", and "did not enjoy" rating scale to allow participants to record their reaction to the interaction when relevant. The back of each daily sheet included an area for (1) a daily rating and (2) free-form feedback surrounding their experience. This is shown in Figure 3.6.

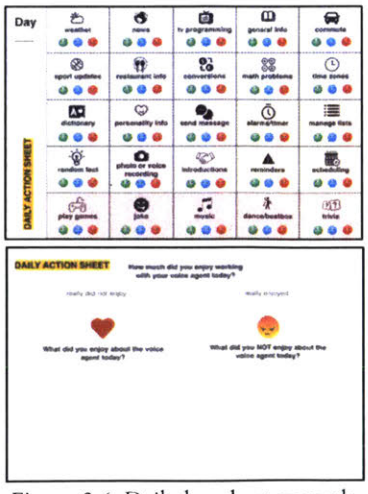


Figure 3.6: Daily log sheet example

While the daily log book aimed to record insights on people's experience with the agent, *an agent wish jar* (Figure 3.7) was used to capture how they wanted the technology to evolve. Participants were given a set of wooden "tokens" to write down wishes and throw into the "wish jar". The simple metaphor of a wish jar was used to equalize differences in comprehension between children, adults, and older adults and allow for a shared probe.

Finally, multiple channels for communication were setup including text messaging, email, and calls to enable families to share free-form photos, videos, notes, and feedback throughout the process. Prompts to send these artifacts were interspersed throughout the participant log book at relevant moments (i.e. agent unboxing). In addition, different colored tokens and pens were provided as a means to create greater ownership during the recording process when multiple participants were in a household.

Figure 3.7: Wish jar and tokens used to capture participants' desires for voice-based agent technology in their home.

STRUCTURED IDEATION CARDS

While participants were exposed to a range of agents as technology probes, another important part of this work was to encourage, support, and capture their ideas of a "dream agent". Prior work in ideation has emphasized the difficulties of the process with regards to "achieving both quantity and diversity of ideas, as well as ideas that are creative and will eventually lead to something that works" (Daly, Seifert, Yilmaz & Gonzalez, 2016). To combat such challenges, design tools and techniques such as design heuristics and inspiration cards have frequently been utilized. Design heuristics refer to "prompts" that allow people to explore design solution spaces and have been shown to be effective in a range of settings including classrooms and engineering teams (Daly, Seifert, Yilmaz & Gonzalez, 2016). Tools such as inspiration cards have also been utilized as aids in the ideation process to provide structure and support shifts in focus (Halskov & Dalsgård, 2006). During early explorations in this work, drawings and "bags of stuff" were used to capture notions of a dream agent (Figure 3.8). While this approach was effective as part of initial explorations, it became increasingly important to create more structure around the ideation. To do so, design heuristics and inspiration cards were used as a basis to develop a set of *agent design cards* as detailed in Figure 3.9.

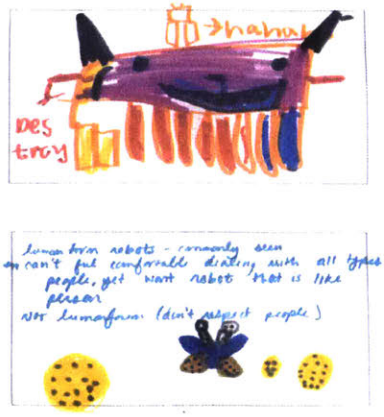


Figure 3.9: Drawings and prototypes of "dream agents" created by children and adults.

Figure 3.9: Agent design card tools to enable participants to go through a structured ideation process for designing their designing the form, function, behavior, and boundaries of their dream agent.

The *agent design cards* were designed to enable people to explore and select eight aspects that are integral to the development of a voice-based agent including the (1 form, (2) materiality, (3) inputs, (4) outputs, (5) personality, (6) location, (7) interaction with other devices in their home and (8) ethics. Accessibility and modularity were key principles that informed the design and development of these cards. To make this process accessible to three generations of participants and equalize differences in understanding of the technology, easy-to-understand language was used. Secondly, to allow facilitators to adapt to the tool based on varying contexts, the cards were designed with modularity in mind. For example, in studies when participants were present in the lab, smaller versions of the cards were used and attached to material artifacts (e.g. cameras, microphones, different materials) to provide additional context (Figure 3.10). Conversely, when sessions were conducted at assisted-living facilities or at homes, these cards could be carried and used as a book.

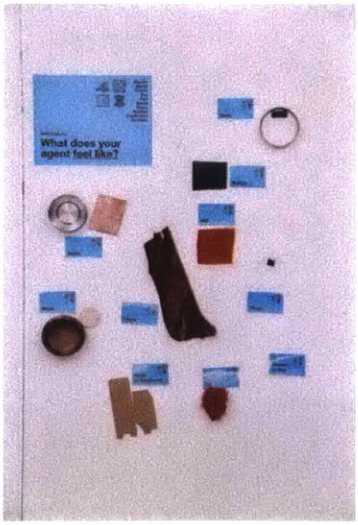
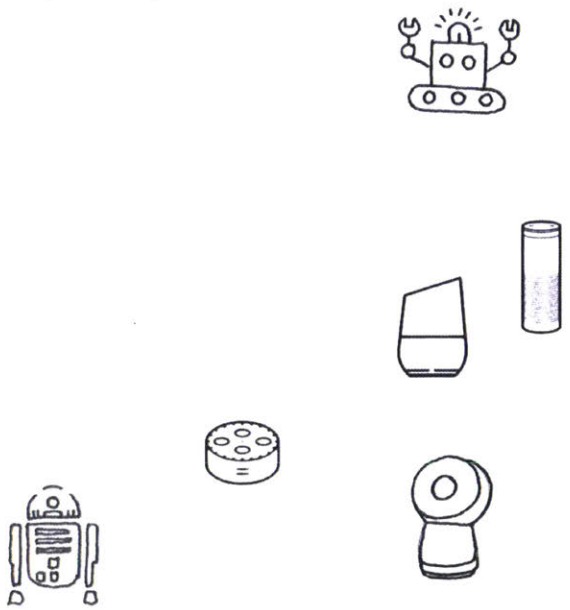


Figure 3.10: Example of cards attached to material artifacts to provide a more tactile and engaging experience.

Part II
First Impressions: Interaction & Discovery

The first study focused on exposing participants to voice-based agent technology and scaffolding their discovery to understand different generations perspectives, desires, and boundaries for the technology. Part II of the thesis describes this process and the resulting insights from these first encounters and impressions. In Chapter 4.0, study objectives, participants, methods, and collected data are described. Chapter 5.0 presents a deep dive into qualitative and quantitative analysis along with discussions on the resulting implications on the design of voice-based agents.



PART II

4.0 Study 1 Overview

What are people's preferences, desires and boundaries for voice-based agent technology in the home?

The core objective of the first study was to focusing on identifying trends in different generations' preferences, desires, and boundaries for voice-based agents. In this section, participant profiles and recruitment is described along with methods and tools utilized in this work.

4.1 Participants

A critical aspect of this work involved engaging a diverse cross-section of the population. To do so, three core populations were recruited including children, adults, and older adults from Cambridge, Boston and Somerville (Massachusetts), Providence (Rhode Island) and Severna Park (Maryland) (Figure 4.1). There was an effort made to represent a range of profiles within each population group with respect to age, gender, living style, housing, education, income (Table 4.1). Data presented in this section was captured from 48 participants (12 children, 12 adults, and 24 older adults) ranging from ages 5 to 98 years old, with a mean age of 42 years old. A histogram of participants' ages is shown in Figure 4.2.

Older adults were recruited through targeted outreach to (1) assisted living centers, (2) senior community centers, and (3) government housing developments, and (4) state elder services agencies. Participants were spread among four key profiles to increase diversity of living styles and experience with technology. The first group of participants resided at an assisted-living center, most having at least a bachelor's degree and

Figure 4.1: Photographs of several participants from three generations recruited for the study.

Feature	Population Statistics
Classification	Children (n=12), Adults (n=12), Older Adults (n=24)
Age	Participants ranged in age from 5-98 years old. The mean age was 49, median was 57.5 with a SD of 28 years.
Gender	Participants were 31% Male (n=15) and 69% Female (n=33).
Living Situation	6% lived with a significant other (n=3), 38% of participants lived alone (n=18), and 56% lived with a family (n=27). Households had 2.6 people on average.
Housing Type	6% lived in a dormitory (n=3), 8% of participants lived in assisted living (n=4), 35% lived in a detached singly-family home (n=17), and 46% lived in a duplex or apartment (n=22). 2 participants did not provide data.
City Type	4% of participants lived in a rural area (n=2), 17% lived in a suburban area (n=8), and 79% lived in an urban area (n=38).
Education	Amongst children, 8% in middle school (n=1), and 92% were in elementary school (n=11). Amongst adults and older adults, the highest level of education completed by participants was 3% middle school (n=1), 8% college associates degree (n=3), 17% high school (n=6), 31% college bachelors degree (n=11) and 42% graduate school (n=15).
Race & Ethnicity	6% identified as White and Hispanic (n=3), 8% of participants identified as White and Asian (n=4), 13% of participants identified as African American (n=6), 13% identified as Asian (n=6), and 58% identified as White (n=28). 1 participant did not provide data.
Income	2% of participants had an income of $125k-$150k (n=1), 6% had an income of $25k-$50k (n=3), 8% had an income of $75k-$100k (n=4), 15% had an income of $50k-$75k (n=7), 15% had an income of $100k-$125k (n=7), and 27% had an income of $0k-$25k (n=13). 7 participants did not report their income.

Table 4.1: Table of demographic breakdowns of overall participant population.

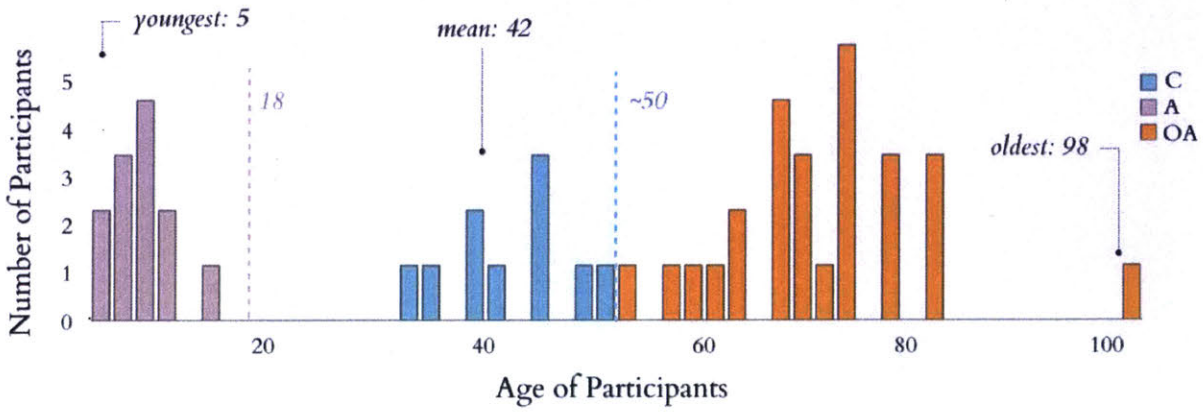


Figure 4.2: Histogram of participant ages and corresponding classifications. The youngest age in the study was 5 years old, while the oldest was 98 years old. The mean age was 42.

high and low income. While the independent high income group had high familiarity with operating technologies such as a smartphone. The next two groups lived independently and varied based on familiarity with technology and generally higher education, participants in the independent low income group often did not own a computer or phone and on average possessed a high-school education. The final group involved older adults living with families. None of the older adults currently owned a voice-based agent in their home. Older adults ranged in age from 50 years old to 98 years old, consistent with other studies in the domain (Vandemeulebroucke, Dierckx de Casterle & Gastmans, 2017). There was a significantly higher proportion of women (n=23) than men (n=4) in this population, which is representative of the gender distribution in the community centers visited and previous technology studies with older adults (Kachouie et al., 2014).

Adults and children were recruited through (1) a mailing list of families in the surrounding area and (2) word of mouth. While recruiting adults and children, an effort was made to balance participants who currently owned a voice-based agent for the home (n=7, 100% Alexa, 14% Google Home) with those who did not (n=41). Genders were well balanced amongst the adult and child participants. Among adults, 42% of participants identified as female (n=5) and 58% identified as male (n=7). Among children, 58% identified as female (n=7) and 42% identified as male (n=5).

4.2 Methods & tools

The first study was divided into four distinct activities including (1) natural interaction to voice-based agents, (2) agent action selection, (3) agent behavior selection, and (4) unstructured "dream agent" design. Sessions were conducted at the MIT Media Lab or at participants' homes depending on their preference between May 2018 – Jan 2018. All sessions were recorded using GoPro cameras and high quality microphones. A photograph of the study setup is shown in Figure 4.3.

Figure 4.3: Photo of study room setup with two card sorting activities for parent and child.

To begin the session, participants were introduced to several voice-based agents including Google Home, Amazon Alexa, and Jibo with example interactions. If participants had no prior experience interacting with voice-based technologies, the

definition provided in the previous chapter was used to frame the technology. After this introduction, participants were given ~20-30 minutes to interact freely with any of the agents. A "bingo card" with examples of actions was provided in the room to help prompt initial queries. Once participants completed this activity, the facilitator prompted them to reflect on their experience interacting with technology through voice and asked questions about the three agents presented. After each session, the location of each agent on the table was randomized in order to mitigate any effects of proximity on agent preference.

After this activity, participants took part in a card sorting activity using the *agent action cards* to identify their preferences and desires for the voice-based agents. One category at a time, participants were presented with a deck of cards and asked to categorize them based on their preference of an agent performing the action. Facilitators began the workshop introducing categories with similarity to current technology affordances first in order to introduce the idea of voice-based agents in accessible terms before bridging into more complex agent actions. For example, the category regarding an agent providing information such as the weather is an action similar to capabilities offered by current phones or iPad. Therefore, this category was postulated to be more easily conceived for a voice-based agent. These categories progressed to more abstract agent actions such as an agent sharing a photo or someone sending a photo through the agent. The participants placed the cards accordingly under three categories on a table: "Yes" if they wanted the social agent action, "Neutral" if they had no preference, and "No" if they did not want the social agent action (Figure 4.3). Throughout the activity, researchers probed with questions, seeking to understand participants' categorization. When children and parents participated as a family together, a divider was placed between their boards to mitigate any influence from the parent's preferences. Additionally, in these scenarios, the child and adult were asked to alternate reflecting on their choices.

Upon completion of the agent action activity selection, the actions the participant categorized as a "Yes" were reintroduced to discuss how an agent could engage with the user (agent behavior activity). Two core concepts were introduced to participants. The agent could be *proactive* and initiate the

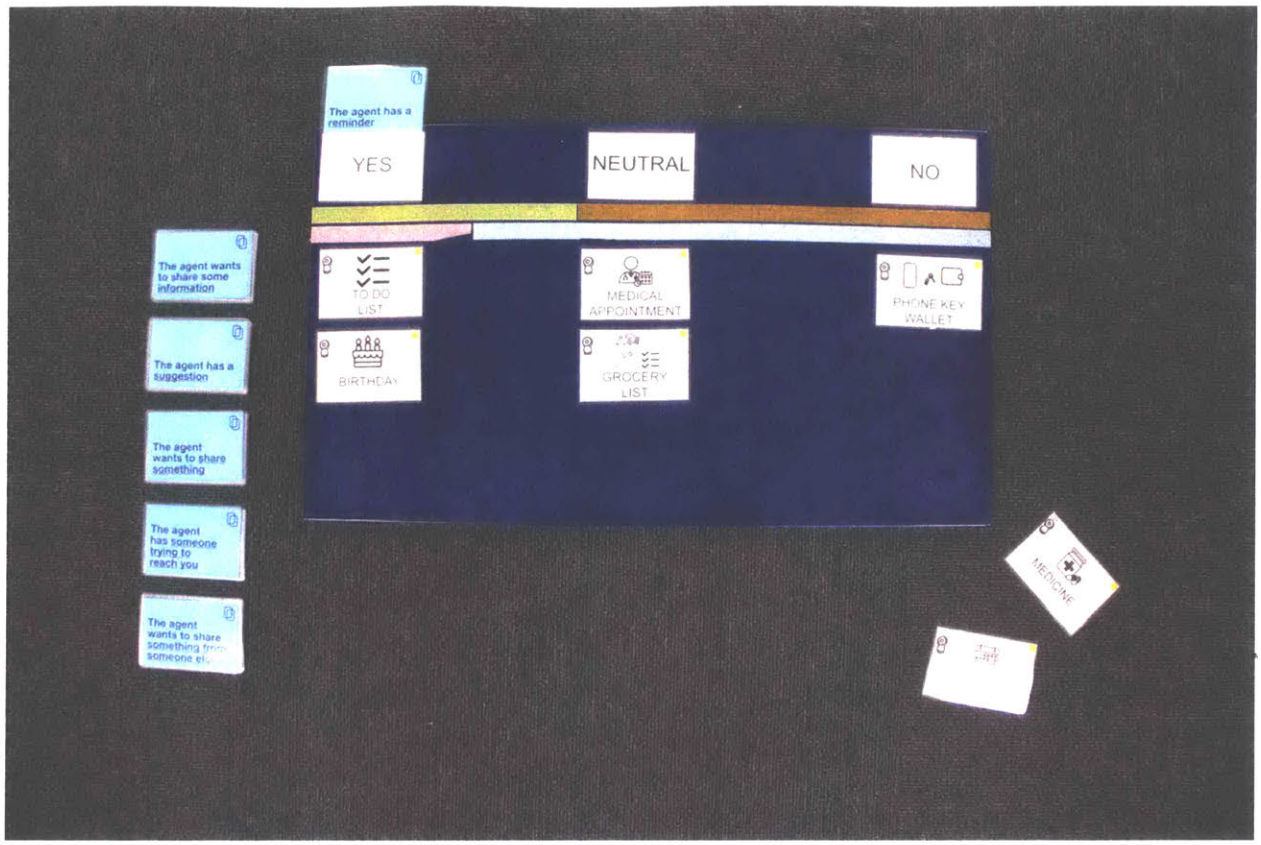


Figure 3.14: Example card sorting of first agent action category. Participants placed cards in "yes", "neutral" and "no" depending on their preference for the agent performing this actions in their life.

interaction, or the user could start the interaction and the agent be *reactive* (Cramer et al., 2009). To contextualize this further, participants were given examples of how this could manifest. After introduction to these concepts, participants were asked to place tokens on any action cards where they wanted the agent to be proactive and take initiative. Researchers further explored their selections, attempting to gain an understanding of how the participants viewed the agent interacting in these ways. For younger participants who had difficulty grasping this notion, an "agent design station" was setup for prototyping their dream agent.

The social agent action and interaction method activities established a common "language of engagement" (Ehn and Sjögren, 1991) that participants could use to further elaborate on ideas surrounding agents. Using this language and these ideas, in the final activity, participants elaborated upon their idea of what a "dream agent" would do in their lives. Participants were provided with worksheets and prototyping tools to allow for unstructured ideation. Facilitators took notes to demonstrate their engagement and valuement of the participants' ideas.

4.3 Data collection and analysis

Over the course of the study, a range of qualitative and quantitative data was collected including (1) audio and video, (2) agent action card choices and (3) demographic data.

Audio and video recordings from all sessions were transcribed using Amazon Mechanical Turk and Rev. The initial read through of these transcribed interviews focused on identifying themes that impacted participants' preferences, desires, and boundaries for voice-based agents. Transcripts from study 1 were then analyzed to identify and organize these themes into categories using an inductive approach (Patton, 2001; Creswell, 2013). An additional two passes were conducted to collect references relevant to the identified themes and further refine and finalize the coding scheme.

The final hierarchical coding scheme contained 33 themes and subthemes organized under 5 distinct theme categories relating to (1) functions the agent could provide, (2) the impact of having the agent, (3) security and privacy, (4) the perceived role of the agent in the household and (5) the capabilities and use of the technology (Table 4.2). A second researcher also qualitatively coded approximately 1/3 of all transcripts from study 1 to check for inter-rater reliability. A Fleiss Kappa of 0.68 for study 1 was determined, indicating a substantial level of agreement (Landis & Koch 2016). Factor prevalence was computed across all participants and by population type in order to identify the most prominent themes. In addition, prevalence was broken down into positive and negative sentiment to allow for further characterization of participants' perspectives.

In addition to qualitative data, a range of quantitative data were collected. Agent action card choices were recorded with each card captured as yes, neutral, or no (2, 1, 0 respectively) to enable analysis on participants' preferences and desires. Data surrounding proactivity was also encoded for each agent action card as a binary value (1 = proactive, 0 = reactive). Demographic data for each participant including age, gender, education, housing, city type, race, ethnicity, people in household, current use of voice-based agents and more was also collected to supplement analysis. A complete breakdown of all data collected in study 1 is provided in Table 4.3 below.

Category	Theme	Sub Theme	Description
Functions agent can help you with	*Health & Wellness*		Does the agent help with reminders surrounding medical appointments & medicine? Does it/could it provide encouragement or reminders to be healthy through eating and exercise habits?
	Social Environment		Is the agent used or aiding to connect with family and friends? Does it become a piece of engagement for family and friends? How does it impact interaction with people?
	Daily Routines	Reminders Information Scheduling/Lists	Does the technology assist you with reminders for things? Does the technology provide you information? Does it help you with your scheduling and list making?
	Novelty & Surprise		Were there any elements of surprise with the technology? Did it provide pleasure or satisfaction because of the novelty? Did it provide breaks in your routine related to the technology's novelty?
Impact	*Laziness*		Does it encourage to not be lazy through movement or exercise? Or does it make you lazier because you can talk to your technology? Does it promote laziness because can use voice to interact with which does not require movement?
	Motivation		Does it encourage you to complete tasks? Could it motivate you in your goals?
	Accountability	For self For others	Does it help you stay accountable to the tasks and/or goals you need to do? Could it hold others in your family accountable?
Security & Privacy	*Security & Privacy*	Data Collection Interaction Experience	Were you worried about data collection? Does it feel comfortable interacting with you? Did you have a concerns around security & privacy with the device? Did you feel concerned having it in your home?
Role	*Perceived Relationship*	Social Companion Assistant/Friend Family/Friend Pet Child	What did you see your voice-controlled agent as being? What role did you see them having?
	Location		Where was it in your home? Why did you put it there?
	Presence in the Home		Was it loud? Soft? Was it randomly interacting? Consistently interacting? How does the form influence space? Take up too much? Takes up little?

Category	Code	Sub Codes	Description
Role	*Proactive Behavior*	Feelings Towards	What did you see your voice-controlled agent as being?
		What?	The actual task or goal
		When?	Appropriateness, urgency
		How?	Notification mechanism, interaction experience
		Who?	Person mediated, choice of people who can do this
	Autonomy	Person Agent	Does the agent have too much autonomy? Do you feel like you have enough control of the technology?
	Attributing Human Qualities		Did it have a personality? What was the personality? How did you feel about the personality?
Technology Capabilities & Use	*Attractiveness of Technology*	Need-to-Have Nice-to-Have	Do you feel like you need the technology and/or its capabilities? Or are they a nice to have?
	Consolidation of Technology		Did you see it as a method to consolidate your technology? Is it a stand-alone piece of technology?
	Ease of Technology		Was it easier to use than other devices? Harder than other devices? Was voice an enjoyable interface? Was is it a frustrating interface?
	Personalization		Can it be personalized to fit needs and preferences? Do you want it to be personalized just to you?

Table 4.2: Coding scheme used for transcripts of sessions from study 1 and study 2. The scheme contains 5 categories with 33 codes.

Data Source	Description of Data	Analysis	QN	QL
Audio & Video	All sessions were recorded using multiple Go Pro cameras to capture different views of the room and audio of interactions with participants.	Sessions were transcribed and qualitatively coded for themes. Prevalence trends were evaluated to generate qualitative findings.		
Agent Action Cards	Agent action card choices were logged as yes, neutral, no (2, 1, 0). In addition, proactivity choices were recorded as binary variables (1 for proactive, 0 for reactive).	Used to explore differences across populations in preferences of agent actions.		
Demographics	Age, gender, income, housing information, people in household, job status, race, ethnicity, education and more.			

Table 4.3: Breakdown of data collected in study 1. QN refers to quantitative while QL refers to quantitative data.

PART II

5.0 Analysis & Discussion

This section presents results from the first study, focusing on identifying trends different generations' preferences, desires, and boundaries for voice-based agents. In this section, (1) first impressions and interactions are described, (2) quantitative data surrounding agent action card choices is explored to identify trends across generations, action categories, and build personas and (2) qualitative data from discussion surrounding these actions is analyzed to identify the drivers for participant choices as well as relevant themes and boundaries for the technology.

5.1 First impressions

Initial interactions with voice-based agents took a range of forms. All generations had difficulty and frustration with remembering to utilize the "wake word" after each interaction, and often continued speaking naturally. This was particularly salient amongst children and older adults.

Children who interacted with Amazon Alexa, Google Home, and Jibo tended to gravitate towards Alexa (Echo Dot) and Jibo after initial interactions. While further exploration is required, this could be due to reasons such as (1) Alexa and Jibo both having "names" and (2) the more overt expression of the light ring on Alexa and Jibo's interactive nature. Children were also more likely to test the same interaction on all three agents to identify their "favorite". When prompted to discuss their preferences for the agents, children were far more likely than adults to describe a distributed system of robots (*"I want Alexa for music, Jibo for dancing, and Google for information"*—child). This was echoed in their descriptions and drawings of their "dream agent" where many children drew a system of many robots,

each with specific capabilities.

Adults in the study were the most comfortable and familiar with the query-answer interaction of current voice-based agents. Most adults minimally explored the agent's capabilities, often trying simple and expected queries such as music and weather. There was limited convergence in adults' initial preference for agents. When asked to describe their dream agent, adults often preferred a consolidation, functional approach—tacking on all capabilities they desired into single form.

Older adults, in many ways, had the widest spectrum of expectations for the agents. This generation struggled most with deciding "how intelligent" the agents were, often expecting them to have natural conversation. Other older adults, who had seen more industrial robots, were initially confused about the objectives of voice-based agents for the domestic environment (*"I'm not sure what to ask now..."*—older adult). Older adults tended to have greater variability in their impressions of Jibo and Alexa than any other generation. Some older adults found it natural and fun to interact with Jibo, with some even exhibiting self-disclosure (*"I'll tell you what its like to live in assisted living. Would you like to hear that? It's pretty lousy. It is no fun to get old. And not to be able to do the things you used to do. I'm stuck inside and you're dependent on other people. My grandson has more independence than I do and he's only 15. I'm telling you, Jibo stay young. Do everything you can do while you're young. Do you like that advice?"*—older adult sharing story with Jibo). Others felt a sense of guilt in not knowing what to ask the agent. This guilt was also present with Alexa, but to a lesser degree. Many older adults also shared that they felt proud and "*hip*" for having interacted with an agent.

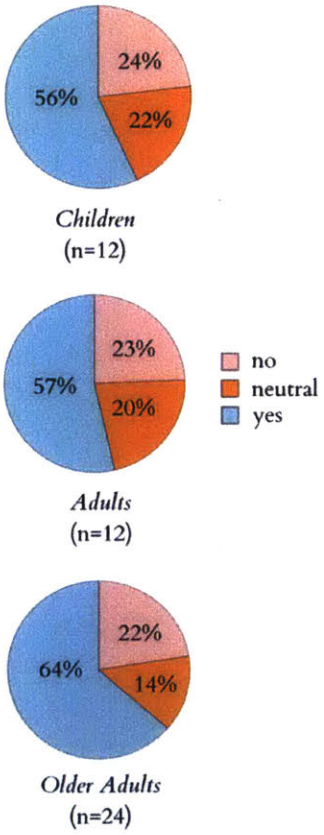


Figure 5.1: Pie charts show breakdown of proportion of 41 agent action cards sorted as no, neutral, and yes by generation

5.2 Quantitative Analysis

AFFINITY TOWARDS AGENTS

Across all generations, over half of the 41 agent action cards were categorized as "Yes" (59% Yes, 17% Neutral, 23% No), indicating a common desire for voice-based agents across all generations. Figure 5.1 and Table 5.1 illustrate the proportion and number of cards sorted into each of the no, neutral, and yes buckets respectively. While all three generations had largely similar distributions across buckets, it is relevant to note that

older adults on average placed a higher number of cards in "yes" and fewer in "neutral" in comparison with adults and children. This effect was determined to be statistically significant for "neutral" cards by a one-way ANOVA ($F(2, 38) = 3.54$, $p = 0.04^*$). This result may be indicative of a marginally higher openness for voice-based agent technology by older adults, though further exploration is required.

	Number of Cards (total cards = 41)		
Generation	No	Neutral	Yes
Children	$\mu = 10.36$, $\sigma = 5.06$	$\mu = 8.91$, $\sigma = 4.06$	$\mu = 23.3$, $\sigma = 7.24$
Adults	$\mu = 10.00$, $\sigma = 7.36$	$\mu = 10.80$, $\sigma = 4.57$	$\mu = 22.00$, $\sigma = 6.61$
Older Adults	$\mu = 11.68$, $\sigma = 8.87$	$\mu = 5.91$, $\sigma = 4.50$	$\mu = 26.08$, $\sigma = 8.92$
ANOVA (across generations)	$F(2, 39) = 0.21$ $p = 0.81$	$F(2, 38) = 3.54$ $p = 0.04^*$	$F(2, 45) = 1.17$ $p = 0.31$

Table 5.1: Mean and standard deviations of number of cards chosen as yes, neutral, and no by generation. Results from ANOVA conducted by bucket across generations are shown.

CHARACTERIZING OVERALL PREFERENCES & DESIRES

In order to understand how participants' preferences and desires varied across categories of agent action cards, breakdowns of the proportion of cards per bucket (yes, no, neutral) and agent action category were computed (Table 5.2).

In addition, to account for participants' placement of cards into the "neutral" bucket, a *weighted preference metric (wpm)* was defined and computed for each participant and associated agent action category. This metric was computed by weighting each yes, neutral, and no card choice by 2, 1, and 0 respectively and normalizing by the maximum possible value for each agent action card category (2*number of cards in the category) as shown below.

$$wpm = \frac{2*(cards_{yes}) + 1*(cards_{neutral}) + 0*(cards_{no})}{2*(cards_{yes} + cards_{neutral} + cards_{no})}$$

A box plot of this metric for each agent action category is shown in Figure 5.2. Normality of the weighted preference metric distribution for each category was evaluated using Shapiro-Wilks. Since not all category distributions were normally distributed, a Kruskal-Wallis H test was conducted to determine

	Proportion of Cards			Percentiles of *wpm*				*wpm*	
Category	No	Neutral	Yes	Min	25th	50th	75th	μ	σ
Reminders	0.19	0.15	0.66	0.14	0.57	0.79	0.96	0.74	0.22
Information	0.19	0.18	0.63	0.29	0.57	0.79	0.93	0.72	0.22
Suggestions	0.31	0.20	0.49	0.00	0.45	0.55	0.80	0.59	0.24
Agent sharing something	0.21	0.16	0.63	0.20	0.60	0.80	0.90	0.71	0.24
Someone trying to reach	0.13	0.12	0.75	0.00	0.67	1.00	1.00	0.81	0.27
Someone sharing something	0.26	0.18	0.56	0.00	0.50	0.69	0.88	0.71	0.29

Table 5.2: Proportion of cards in each bucket as well as quartiles for mean weighted preference metric across participants per category.

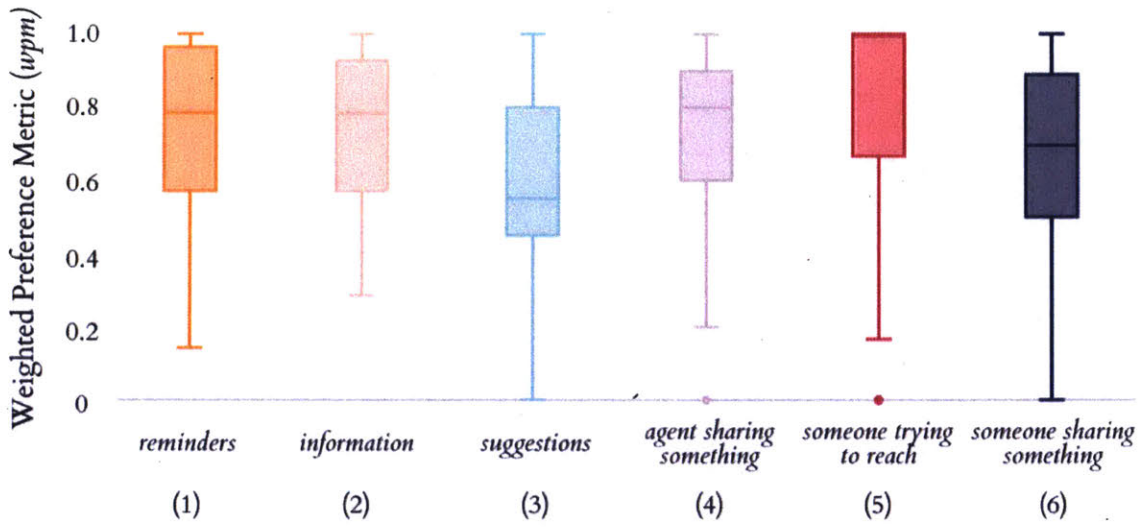


Figure 5.2: Box plot of weighted performance metric distribution per agent action category. Reminders, information, and the agent sharing something have similar distributions. Suggestions

if the participants' preferences varied across different agent action categories. Each category had 48 values for the weighted preference metric, arising from each of the 48 participants. A Kruskal-Wallis H test showed a statistically significant difference in weighted preference between the 6 groups with $H(5) = 27.61$, $p^{*}=4.34e^{-05}$. Corresponding medians and percentiles for each category are shown in Table 4.2. Reminders, information, and the agent sharing something (e.g. playlist, story) were similarly distributed while suggestions, someone trying to reach a person through the agent (e.g. phone call) and someone sharing something through the agent (e.g. photo or message) varied in their distributions from other categories.

To better understand the nature of these differences and evaluate significance, pairwise Mann-Whitney tests were conducted across each of the 6 agent action categories. Results

from this are shown in Table 5.2. A series of Mann-Whitney tests indicated that weighted preference for suggestions was significantly lower (median = 0.55) than reminders (median = 0.79, U=745.00 $p^*=1.39e^{-03}$), information (median = 0.79, U=785.00, $p^*=3.54e^{-03}$), the agent sharing something (median = 0.80, U=811.00, $p^*=6.08e^{-03}$), and someone trying to reach a user through the agent (median = 1.00, U=531.00, $p^*=1.43e^{-06}$). The suggestions category places the agent in the role of helping the user form and maintain habits such as drinking water, meditating, or learning something new. As such, it implies a greater degree of goal-directed behavior and autonomy on the part of the agent in comparison with more functional agent action categories such as reminders, information and the agent sharing something. The suggestions category also appears bimodal in its distribution implying that it may be highly dependent on person-based qualities, though further exploration is required.

Conversely, the agent action category exploring the idea of someone trying to reach a user through the agent had a significantly higher weighted preference (median = 1.00) than other categories including reminders (median = 0.79, U = 836.00 $p^* = 8.53e^{-03}$), information (median = 0.79, U = 799.50, $p^* = 4.24e^{-03}$), the agent sharing something (median = 0.80, U = 768.00, $p^* = 2.02e^{-03}$), suggestions (median = 0.55, U = 531.00, $p^* = 1.43e^{-06}$) and someone sharing something with the user through the agent (median = 0.69, U = 669.50, $p^* = 1.64e^{-04}$).

Table 5.2: Pairwise Mann-Whitney tests across all 6 agent action categories. Category (3) suggestions and (5) someone trying to reach the user different significantly from other categories of agent actions.

	Mann-Whitney Results By Category Pair (n = 48 for each)					
Category	(1)	(2)	(3)	(4)	(5)	(6)
(1) *Reminders*	-	U=1115.00, p=0.39	U=745.00 $p^*=1.39e^{-03}$	U=1151.00 p=0.49	U=836.00 $p^*=8.53e^{-03}$	U=971.50 p=0.09
(2) *Information*	-	-	U=785.00, $p^*=3.54e^{-03}$	U=1131.00, p=0.44	U=799.50, $p^*=4.24e^{-03}$	U=1021.00, p=0.17
(3) *Suggestions*	-	-	-	U=811.00, $p^*=6.08e^{-03}$	U=531.00, $p^*=1.43e^{-06}$	U=947.00, p=0.067
(4) *Agent sharing something*	-	-	-	-	U=768.00, $p^*=2.02e^{-03}$	U=1039.00, p=0.20
(5) *Someone trying to reach*	-	-	-	-	-	U=669.50, $p^*=1.64e^{-04}$
(6) *Someone sharing something*	-	-	-	-	-	-

It is interesting to note that the median weighted preference for the agent acting as a means for others to *reach* the user (e.g. phone call, status update) is significantly higher than the median value for the agent acting as a means for others to *share* something with the user. To better understand this observation, it is valuable to note that these categories differ with respect to the time-sensitivity of the intended interaction while both exploring ideas of technology consolidation and agent-mediated human-human interaction. For further verification of results, Mann-Whitney tests were conducted across category pairs for the *proportion of cards placed in the "yes" bucket* alone without weighting for cards placed in "neutral". This set of tests identified identical category pairs as being statistically significant.

HOW GENERATIONS THINK DIFFERENTLY

As evident from initial results, participants' preferences for agent actions varied across categories. In order to understand the drivers of these observations, it is important to consider whether these differences in preferences are a function of the generation or characteristic of the larger population. Figure 5.3 below shows

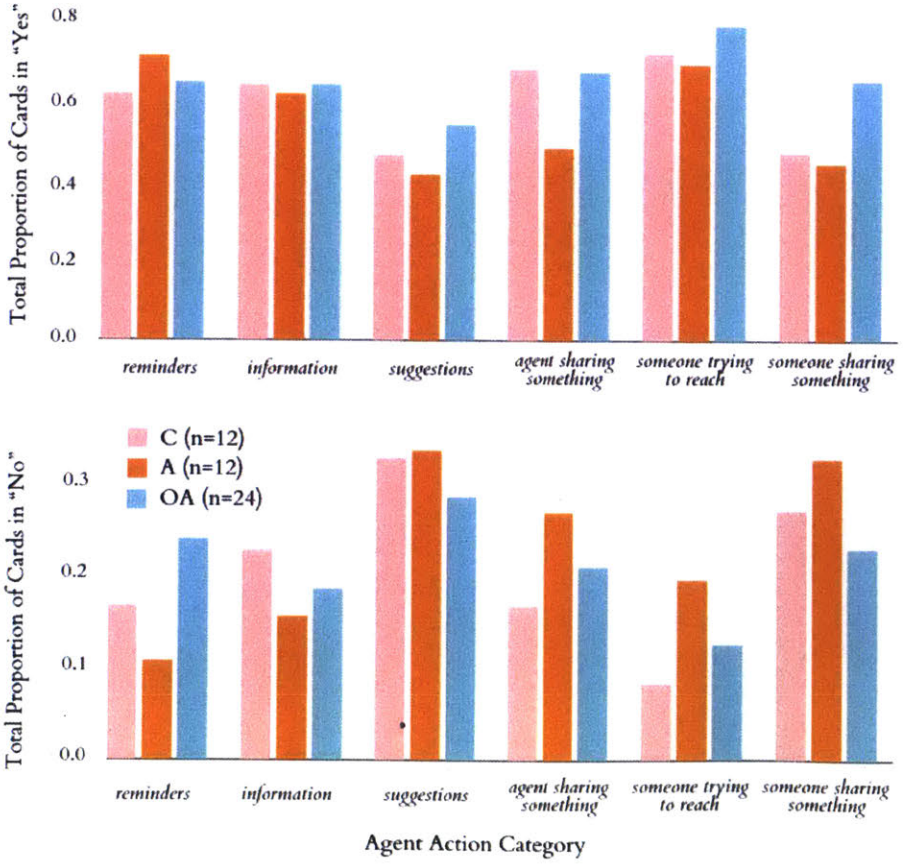


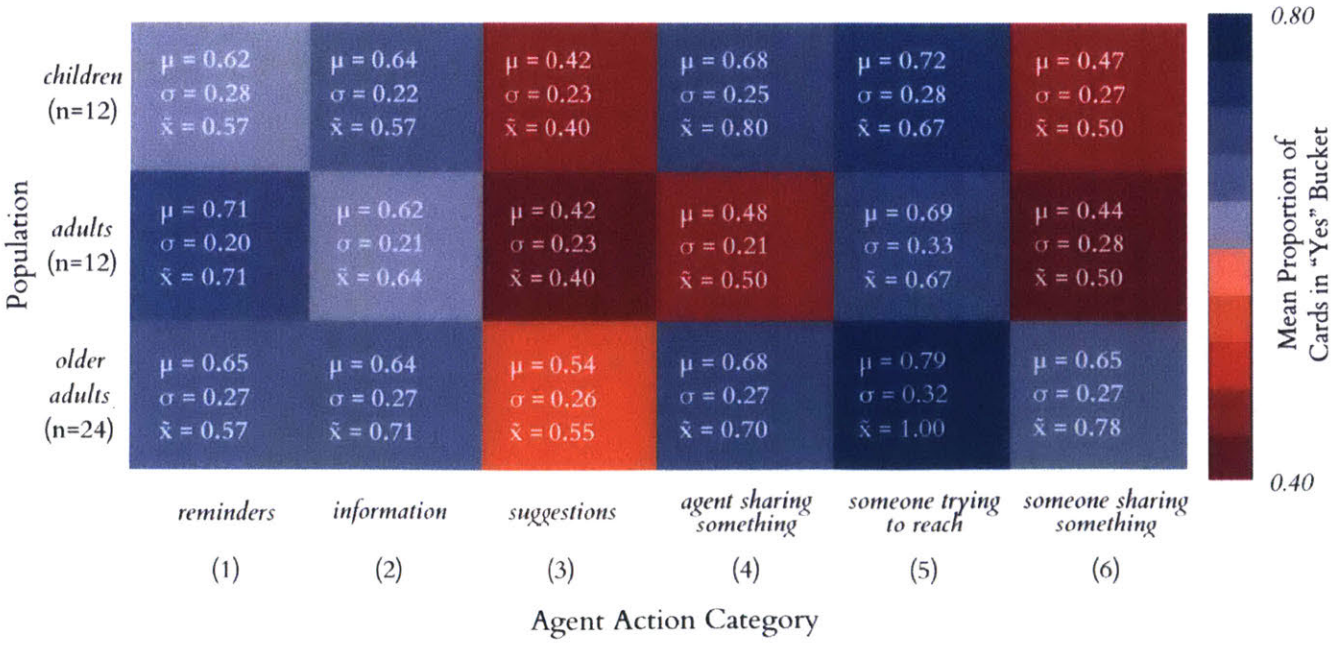
Figure 5.3: Bar charts for total proportion of cards placed in the "Yes" and "No" buckets respectively by generation and agent action category.

bar charts of the total proportion of cards placed in the "Yes" and "No" buckets respectively by category and generation. As evident, categories had differences across generations in placement of cards in the "Yes" and "No" buckets respectively. These differences were strongest for categories such as the agent sharing something and someone sharing something through the agent. In addition, while categories such as reminders and information had reasonable consistency for the proportion of cards placed in the "Yes" bucket, there was greater differences in the proportion of cards placed in the "No" bucket, indicating subtle differences in generational choices.

To further characterize these differences, the mean proportion of cards placed in the "yes" bucket by generation (12 adults, 12 children, 24 older adults) and agent action category was computed, as shown in the heat map in Figure 5.4. The proportion of cards placed in the "yes" bucket was compared without weighting for those placed in "neutral" to better isolate participants' stronger desires for agent actions. As evident from the heat map, categories relating to the agent sharing something with the user (e.g. playlist, story) and someone sharing something with the user through the agent (e.g. message, sending a song) had the greatest variability across generations. Reminders ($\mu_C = 0.62$, $\mu_A = 0.71$, $\mu_{OA} = 0.65$), information ($\mu_C = 0.64$, $\mu_A = 0.62$, $\mu_{OA} = 0.64$), and someone trying to reach the user through the agent ($\mu_C = 0.71$, $\mu_A = 0.69$, $\mu_{OA} = 0.79$), were seen as

s

Figure 5.4: Mean, standard deviations and median for proportion of cards in the "yes" bucket across generations and agent action category.



Category	Across Generations Comparison (Kruskal-Wallis H Test) of Proportion of Cards in "Yes"	
Reminders	H(2) = 0.79	p = 0.67
Information	H(2) = 0.09	p = 0.96
Suggestions	H(2) = 1.79	p = 0.41
Agent sharing something	H(2) = 4.90	p = 0.09
Someone trying to reach	H(2) = 1.55	p = 0.46
Someone sharing something	H(2) = 5.79	p = 0.05*

Table 5.5: Results from series of Kruskal-Wallis H tests in proportion of cards placed in the "yes" bucket across generation groups for each agent action category.

valuable across all generations. Conversely, participants of all generations expressed a lower desire for suggestions from the agent.

A series of Kruskal-Wallis H tests were conducted comparing the proportion of cards placed in the "yes" bucket across children, adults, and older adults for each agent action category to determine statistical significance of these findings (Table 5.5). A Kruskal-Wallis H test showed a statistically significant difference in agent action preference between the 3 generations for someone sharing something with the user through the agent ($H(2) = 5.79$, $p^* = 0.05$). A low, but weak significance ($p = 0.09$) was also found to be present for the agent sharing something with the user (e.g. playlist, story).

To identify which generations were statistically different from each other in these two categories, pairwise Mann-Whitney tests were conducted for each generation pair (Table 5.6). Results indicated that preference for people sharing something through the agent (e.g. messages, sharing a photo or video) was significantly lower for children (median = 0.5) and adults (median = 0.5) than older adults (median = 0.78). In addition, preference for the agent sharing something (e.g. playlist, story) was found to be significantly lower for adults

Category	Mann-Whitney Results By Generation Pair of Proportion of Cards in "Yes"		
	Children-Adults	*Children-Older Adults*	*Adults-Older Adults*
Someone sharing something	U=64.50, p=0.34	U=85.00, p^*=0.02	U=87.50 p^*=0.02
Agent sharing something	U=41.50, p^*=0.04	U=142.50, p^*=0.49	U=84.00, p^*=0.03

Table 5.6: Results from series of pairwise Mann-Whitney tests on various generation pairs for categories.

(median = 0.5) than children (median = 0.8) and older adults (median = 0.7). Histograms and distributions of the *weighted proportion metric* by each generation and agent action category are also provided in Figure 5.5 for additional context.

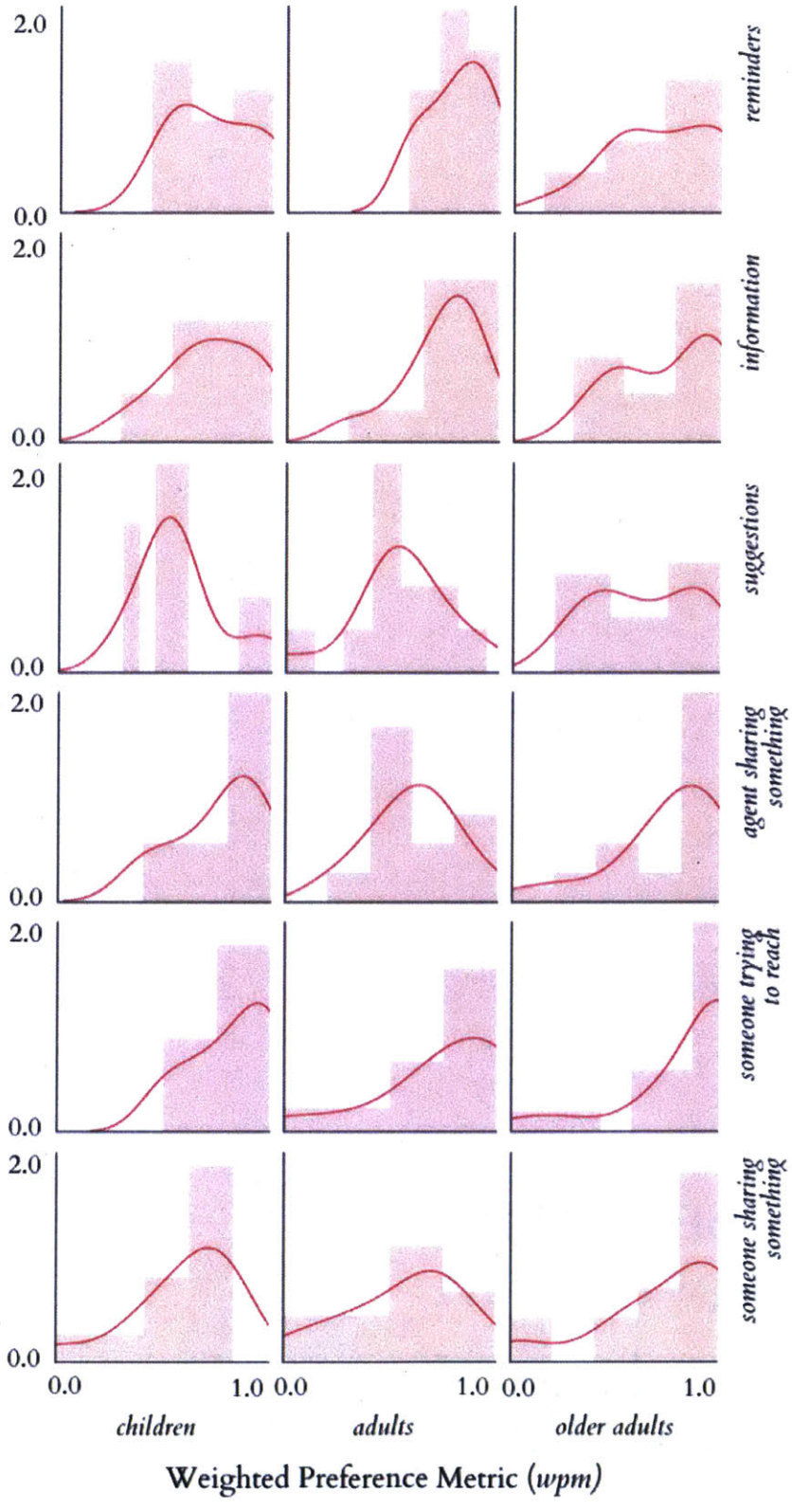


Figure 5.5: Histogram and corresponding distributions by generation and category of the weighted preference metric.

Given these findings, it is relevant to recall that both agent action categories in which the adult population expressed a lower preference embody social elements. In one, the agent itself exhibits social behaviors by sharing music, stories, a thank you and more with the user. In the other, the agent acts as a mediator for social interactions with others through sharing of photos, video, messages and more. Overall distributions of the weighted preference metric across each category and generation also support these findings. Adults exhibited a stronger preference towards more functional agent actions—such as reminders information, and someone trying to reach the user (e.g. phone call) than suggestions from the agent and socially-driven action categories. Although children and older adults expressed a similar preference for the functional agent action categories, they also exhibited higher affinity towards socially-driven action categories than adults. In particular, results showed significantly higher preference towards others using the agents as a means to share social content (e.g. photos, videos, messages) in the older adult population. Older adults and children both expressed a higher preference for the agent itself sharing something (e.g. playlist, story, thank you).

WHAT DIFFERENT GENERATIONS WANT

To dissect these results further, the proportion of participants from each generation who placed each agent action card into the "yes" bucket was computed for each action. These values are shown in the heat map in Figure 4.5 along with radar charts to illustrate the overall desire for agent actions by generation.

Older adults had broader, and stronger desire for agent actions than adults and children. Agent actions including (1) reminders for medical appointments (0.75), meeting someone (0.71), and birthdays (0.75), (2) information about nutrition facts (0.71) and weather (0.71), (3) suggestions for social events (0.75) in the area and learning something new (0.83) and (4) music playlists made by the agent were the most desired. In addition, most older adults were also motivated by the prospect of the voice-based agent acting as a means to consolidate existing technology and communicate with people for phone calls (0.75) and messages (0.81). While older adults as a population were the most open to voice-based agent actions, there was strong opposition to the idea of the agent providing suggestions around

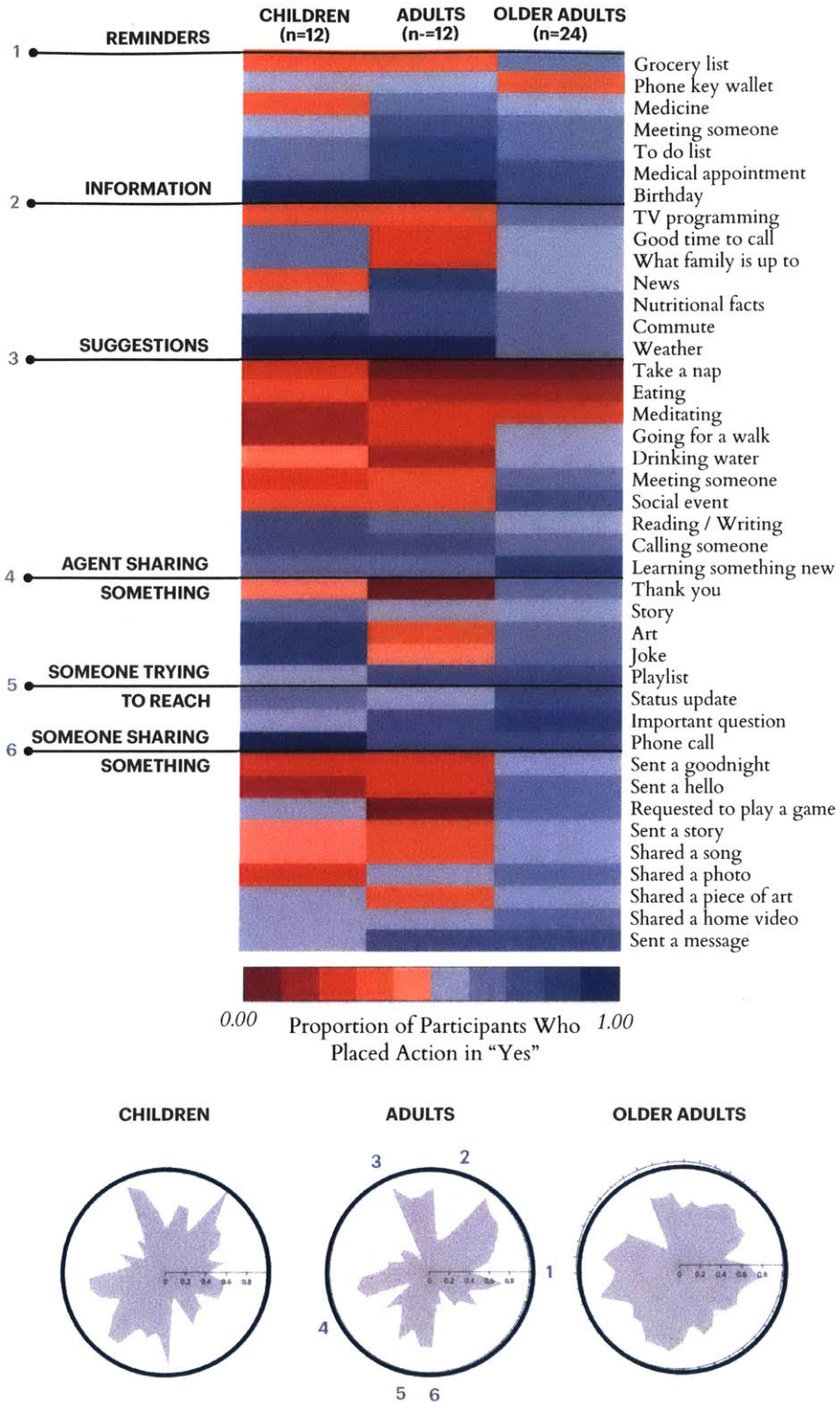


Figure 5.6: Proportion of participants who placed agent action card in the "yes" bucket by card and generation. Radar charts by generation are provided for overall context. Numbers around chart indicate corresponding agent action card category.

basic needs such as taking a nap (0.16) and eating (0.21).

Unlike older adults, children and adults were far more selective in their desire for agent actions. Children often wanted agents to (1) remind them of birthdays (1.00), (2) provide information surrounding commutes (e.g. when the school bus will pickup) (0.83) and weather (e.g. what to wear in the morning) (0.92), (3) give suggestions to read or write (0.75) and call friends or family (0.75), (4) entertain them through pictures (0.83) or jokes (0.83), (4) become a means to call others (0.92). Adults saw the agent primarily as a functional technology, often selecting actions surrounding reminders and information including (1) reminders for meeting someone (0.75), to-do lists (0.83), medical appointments (0.83), and birthdays (0.92), and (2) information about the news (0.83), nutritional facts (0.75), commute (0.75), and weather (0.92). Adults were also open to the agent suggesting they call someone they hadn't spoken to in a while (0.75) and people reaching them through the agent, particularly for time-sensitive needs such as phone calls (0.75), messages (0.75) or questions (0.75). Conversely, both children and adults were often opposed towards the agent providing suggestions for day-to-day habits such as taking a nap (A = 0.08, C = 0.33), eating (A = 0.25, C = 0.42), meditating (A = 0.33, C = 0.25), going for a walk (A = 0.33, C = 0.25), and drinking water (A = 0.25, C = 0.5). The most contentious agent action across the generations was the idea of the agent saying *"thank you"* (e.g. after an interaction or game) (C = 0.5, A = 0.17, OA = 0.67), with adults being the most opposed and older adults the least.

IDENTIFYING PERSONAS

To better explore the drivers for these differences beyond generational factors, a clustering approach was utilized to identify similar groups of users based on agent action card selection patterns. The weighted preference metric was used to capture more of the variability in user selection (Yes = 2, Neutral = 1, No = 1). An elbow function and hierarchical clustering map were utilized to identify the appropriate number of clusters for *k-means* clustering. After several iterations, a cluster size of 5 was chosen for identifying distinct personas in the population. The resulting means, standard deviations, and population breakdowns for these clusters across each of agent action categories are shown in Figure 5.7 and Table 5.7.

Agent Action Category	Mean and SD of *wpm* by Cluster				
	C1: Open to everything (n = 16)	*C2: Agent as tool (n = 3)*	*C3: Agent as social tool (n = 19)*	*C4: No social from others (n = 3)*	*C5: Skeptics (n = 7)*
Reminders	μ = 0.90 σ = 0.13	μ = 0.95 σ = 0.08	μ = 0.62 σ = 0.15	μ = 0.93 σ = 0.12	μ = 0.50 σ = 0.20
Information	μ = 0.92 σ = 0.08	μ = 0.60 σ = 0.15	μ = 0.68 σ = 0.19	μ = 0.69 σ = 0.11	μ = 0.47 σ = 0.21
Suggestions	μ = 0.84 σ = 0.13	μ = 0.50 σ = 0.00	μ = 0.51 σ = 0.11	μ = 0.58 σ = 0.37	μ = 0.31 σ = 0.17
Agent sharing something	μ = 0.88 σ = 0.10	μ = 0.33 σ = 0.12	μ = 0.74 σ = 0.15	μ = 0.57 σ = 0.37	μ = 0.44 σ = 0.27
Someone trying to reach	μ = 0.97 σ = 0.09	μ = 1.00 σ = 0.15	μ = 0.81 σ = 0.15	μ = 1.00 σ = 0.00	μ = 0.29 σ = 0.00
Someone sharing something	μ = 0.86 σ = 0.13	μ = 0.78 σ = 0.18	μ = 0.57 σ = 0.06	μ = 0.04 σ = 0.05	μ = 0.29 σ = 0.25
Population Breakdown	C = 3 A = 3 OA = 10	A = 2 OA = 1	C = 7 A = 3 OA = 9	C = 1 A = 1 OA = 1	C = 1 A = 3 OA = 3

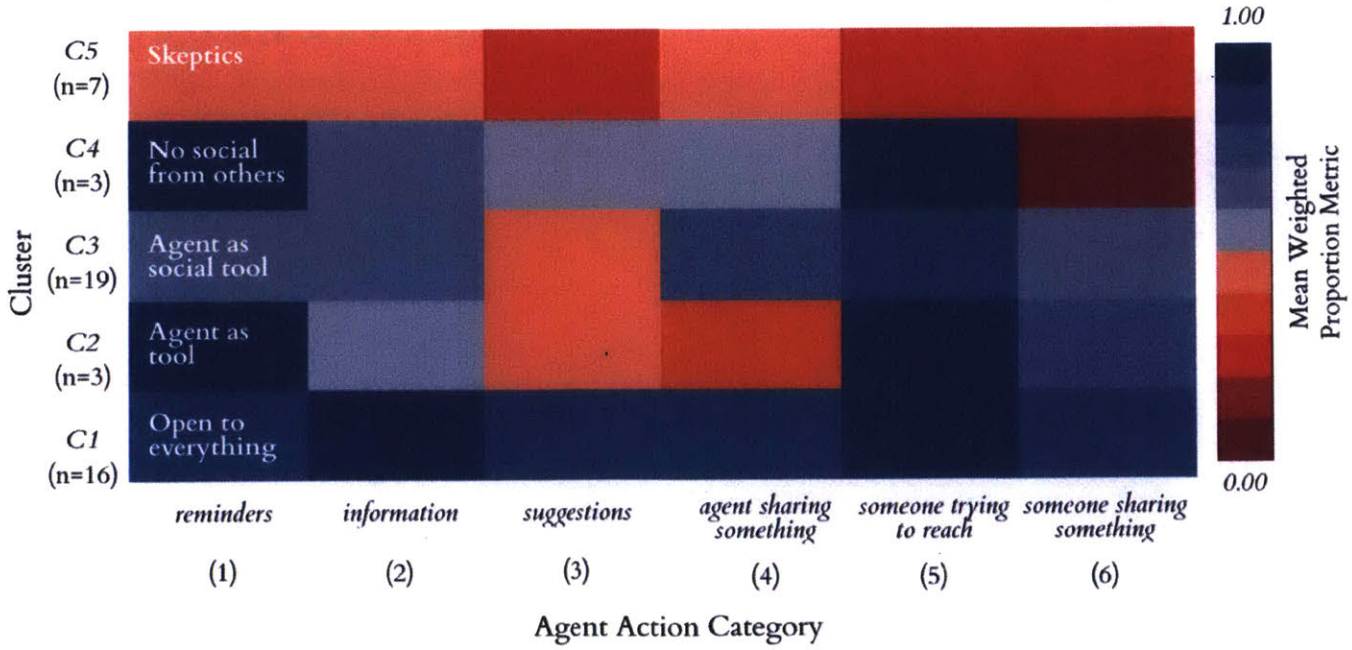


Table 5.8: Breakdown of mean and standard deviation of weighted proportion metric by agent action category and cluster (as determined by *k-means*). Breakdowns of the population in each cluster are also provided. Figure 5.8 shows a corresponding map of the mean weighted proportion metric by agent action category and cluster.

On the basis of these clusters, a series of corresponding personas were defined and supported through qualitative data. These are described in greater detail below:

Cluster	Persona
Cluster 1 ***Open to everything***	This cluster describes users who were very open to exploring a broad range of roles for voice-based agents in their lives. A disproportionate number of participants from the study in this group were older adults.
Cluster 2 ***Agent as a tool***	This cluster described users who were open to voice-based agents as a *tool* for functional and social goals, but were less open to the agent itself expressing goal-directed behavior through suggestions or sharing content (e.g. music, drawing).
Cluster 3 ***Agent as a social tool***	This cluster described users who were also very open to exploring a broad set of roles for voice-based agents (including functional and social), with a slightly higher affinity towards socially-driven actions such as the agent sharing something and people reaching or sharing content with the user through the agent. Participants in this group tended to be either older adults or children.
Cluster 4 ***No social content from others***	This cluster described users who are similar to clusters 2 and 3 in their desire for the voice-based agent acting as a functional tool. In addition, users in this cluster were open to the idea of the agent having goal-directed behavior through suggestions and sharing content. However, they had a strong non-desire for the agent acting as a means for others to share content. This may suggest a repulsion towards the idea of an agent acting as a mediator of human-human connection.
Cluster 5 ***Skeptics***	This cluster described users who were skeptical towards the idea of voice-based agents in their life across all categories. This cluster was primarily composed of adult and older adult participants.

Table 5.9: Developed personas based on results from clustering analysis.

5.3 Qualitative Analysis

To better understand the drivers of these observed differences across people, qualitative data from transcripts was analyzed. Presence values (i.e. proportion of participants who brought up the particular theme at least once) were determined for each theme defined in the coding scheme. An overall breakdown and corresponding chart of the presence of these themes across participants is shown in Table 5.10 and Figure 5.9.

Category	Theme	Subtheme	Proportion of Participants With Theme Present					
			Total (n = 43)	Positive	Negative	C (n=12)	A (n=9)	OA (n=22)
Functions Agent Can Help You With			0.98	0.98	0.58	0.92	1.00	1.00
	Daily Routines		0.95	0.95	0.30	0.83	1.00	1.00
		Reminders	0.79	0.77	0.19	0.58	0.78	0.91
		Information	0.79	0.77	0.07	0.67	0.44	0.91
		Scheduling & Lists	0.40	0.37	0.07	0.16	0.55	0.45
	Social Environment		0.88	0.83	0.35	0.83	0.78	0.95
	Novelty		0.56	0.56	0.14	0.66	0.77	0.41
	Health & Wellness		0.49	0.44	0.05	0.17	0.22	0.77
Technology Capabilities & Use			0.98	0.79	0.89	0.83	1.00	1.00
	Attractiveness of Technology		0.79	0.40	0.72	0.50	1.00	0.86
		Need-to-have	0.74	0.28	0.70	0.42	1.00	0.82
		Nice-to-have	0.28	0.21	0.09	0.08	0.11	0.45
	Consolidation of Technology		0.74	0.58	0.40	0.50	0.67	0.91
	Personalization		0.58	0.51	0.23	0.42	0.89	0.55
	Ease of Technology		0.26	0.19	0.09	0.00	0.00	0.50
Role			0.86	0.67	0.56	0.75	1.00	0.86
	Autonomy		0.77	-	-	0.58	0.89	0.82
		Person	0.69	-	-	0.33	0.89	0.77
		Agent	0.65	-	-	0.42	0.78	0.72
	Proactive Behavior		0.74	0.74	0.49	0.50	0.89	0.82
		What	0.51	0.47	0.21	0.17	0.78	0.59
		When	0.52	0.40	0.16	0.33	0.67	0.50
		How	0.49	0.37	0.28	0.08	0.44	0.68
		Who	0.21	0.21	0.00	0.25	0.33	0.14
		Feelings Towards	0.05	0.00	0.05	0.00	0.00	0.09
	Perceived Relationship		0.53	-	-	0.33	0.67	0.59
		Assistant	0.32	-	-	0.08	0.44	0.41
		Social Companion	0.28	-	-	0.17	0.33	0.32
		Family or Friend	0.12	-	-	0.17	0.00	0.14
		Child	0.02	-	-	0.00	0.00	0.05
		Pet	0.02	-	-	0.00	0.00	0.05
	Attributing Human Qualities		0.42	-	-	0.33	0.44	0.45
	Presence in the Home		0.21	0.07	0.17	0.00	0.22	0.32
	Location		0.02	0.02	0.02	0.00	0.00	0.05

Impact	0.46	0.35	0.21	0.25	0.44	0.59
Motivation	0.37	0.33	0.07	0.25	0.44	0.41
Accountability	0.26	-	-	0.08	0.22	0.36
For Self	0.21	-	-	0.08	0.11	0.32
For Others	0.09	-	-	0.00	0.22	0.09
Laziness	0.21	0.07	0.21	0.00	0.11	0.36
Security & Privacy	0.41	0.14	0.30	0.50	0.55	0.32
Interaction Experience	0.37	0.11	0.26	0.50	0.44	0.27
Data Collection	0.09	0.02	0.07	0.00	0.22	0.09

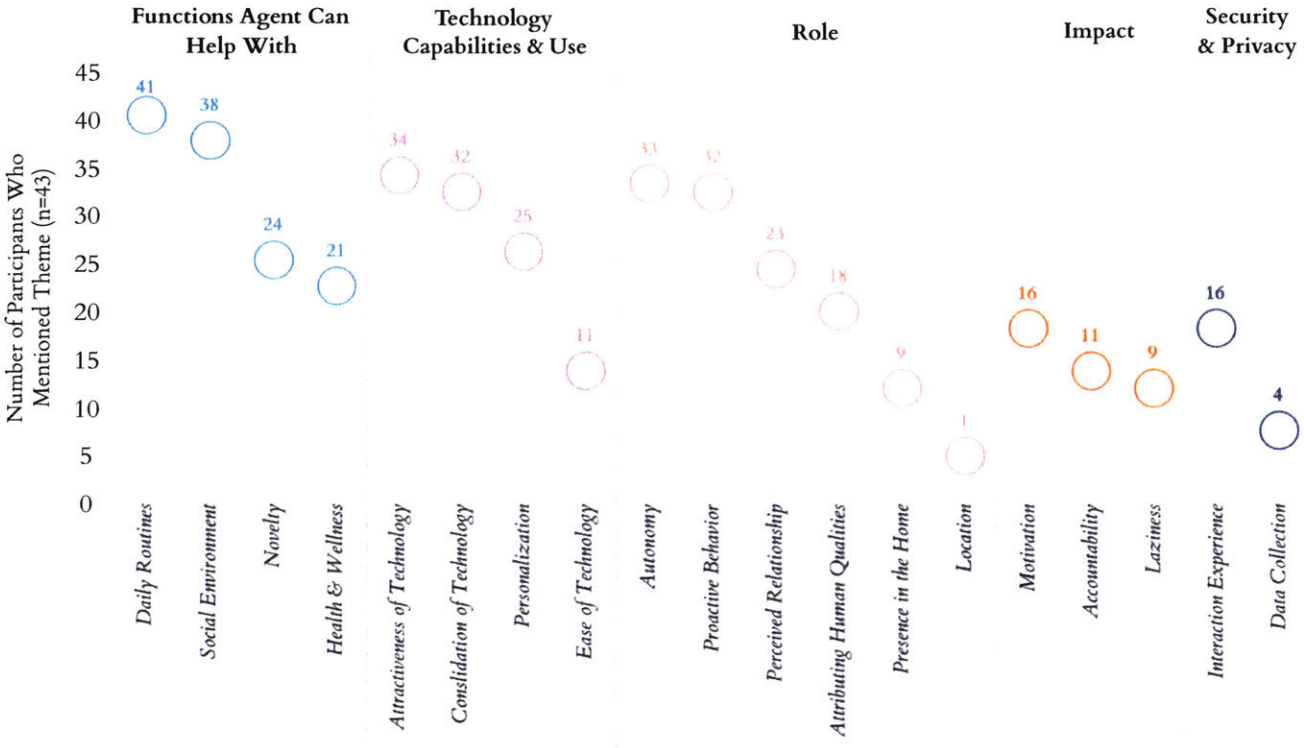


Table 5.10 summarizes the presence of codes by overarching theme category, theme, and associated subthemes when applicable. In addition, the proportion of participants who discussed each code positively and negatively is summarized. These proportions are also broken down by generation (children, adults, older adults) for additional context.

As is evident from these breakdowns, the most common areas of discussion for all participants were surrounding functions the agent could help the user with, the capabilities of the technology, and the potential role the technology could take on.

Table 5.10: Breakdown of presence by category, theme, and subtheme from transcripts by generation. Values represent proportion of participants who spoke about specific theme. Figure 5.9 shows a corresponding chart of number of occurrences (n=42) for each category and theme.

Almost all participants discussed positive aspects of functions the agent could help (0.98), while some also highlighted negative consequences (0.58). Dialogue about these functions focused significantly on the integration of the technology into daily routines and its potential impact on the social environment. With regards to technology capabilities and use, participants focused on the attractiveness (or unattractiveness) of the technology as well as its potential (or non-potential) to consolidate other technologies. Sentiments surrounding technology capabilities and use were rather mixed, with many participants discussing both positive (0.79) and negative (0.89) perspectives. Discussion on the role of the agent focused on aspects of autonomy—both for the individual and the agent—as well as feelings towards proactive behavior. While participants also addressed ideas around technology impact, security and privacy, these were less frequent in comparison with other themes.

While the relative presence of these theme categories across generations was largely consistent, there were key differences in the presence of themes and subthemes. Figures 4.8 (a), (b), and (c) shows treemaps by generation of the proportion of people who discussed each subtheme. With respect to functions the agent could help with, children (0.83) and older adults (0.95) were more likely than adults (0.78) to reference its potential impact on their social environment. This is consistent with findings presented in the quantitative analysis above, showing these generations' greater affinity for social elements in the technology. Older adults (0.77) were also far more likely to discuss health and wellness with regards to the functions the agent could provide than children (0.17) and adults (0.22). Additionally, novelty was more likely to be addressed by adults (0.71) and children (0.66) than older adults (0.41).

Discussions around technology capabilities and use exhibited more distinct differences across generations. No children or adults discussed their perspectives on the ease of using voice-based agent technologies, while half of the older adults found this to be relevant. A significantly higher proportion of older adults (0.91) also discussed ideas around technology consolidation than adults (0.67) and children (0.50). On the other hand, adults were disproportionately more likely to address their perspectives towards personalization (0.89) and technology attractiveness (1.00).

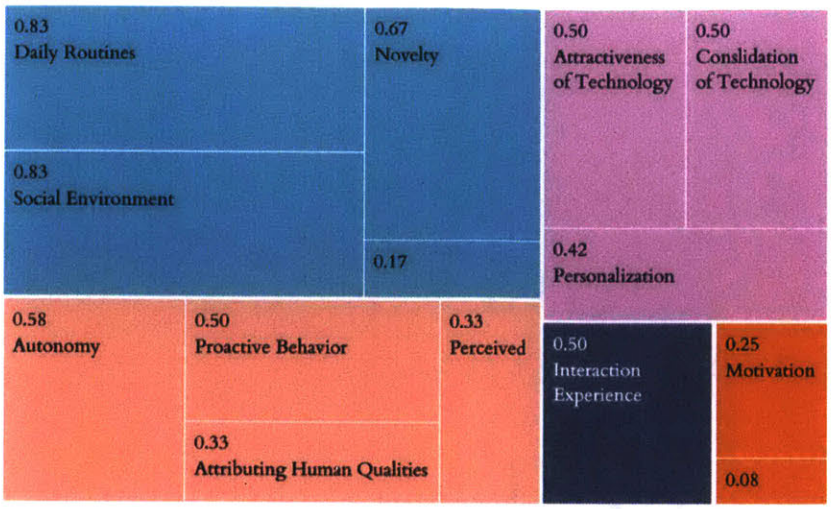


Figure 4.8 (a): Tree map of most relevant subthemes for children. Values shown indicate proportion of children who discussed the theme during card sorting and interviews.

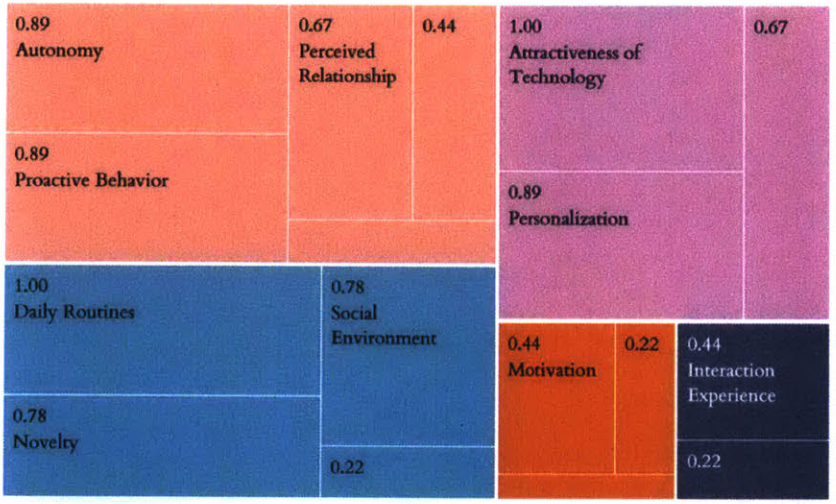


Figure 4.8 (b): Tree map of most relevant subthemes for adult. Values shown indicate proportion of adult who discussed the theme during card sorting and interviews.

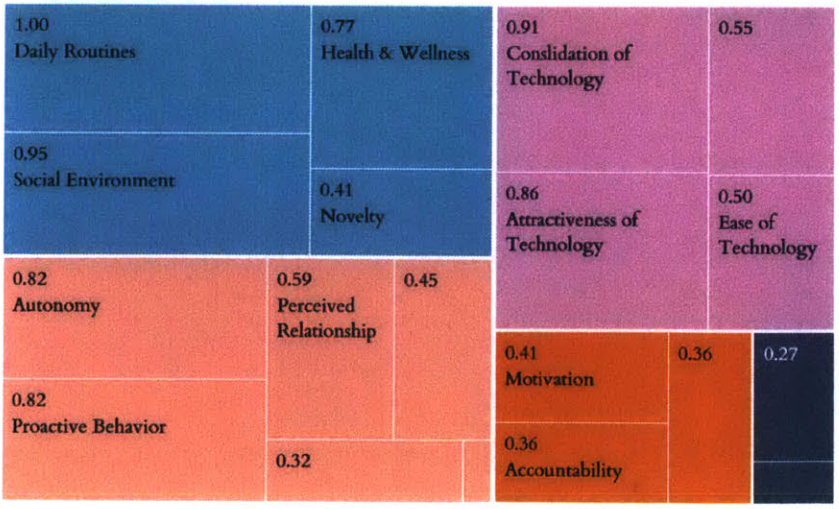


Figure 4.8 (c): Tree map of most relevant subthemes for older adults. Values shown indicate proportion of adult who discussed the theme during card sorting and interviews.

With regards to the role of the agent, all generations discussed ideas around autonomy. However, a much higher proportion of adults (0.86) and older adults (0.77) discussed their own autonomy in the agent-person relationship than children (0.33). Further distinctions also existed in the way that generations described their perceived relationship with the agent. Older adults were the most diverse in their perception, describing the agent as an assistant (0.41), social companion (0.32), family or friend (0.14), child (0.05) or even pet (0.05). Adults on the other hand primarily spoke about the agent as an assistant (0.44) or social companion (0.41). Children, though less likely to discuss notions around agent relationship, described the agent as a social companion (0.17), family or friend (0.17) and sometimes assistant (0.08). These differences are consistent with findings relating to intergenerational agent action preferences from quantitative analysis.

Interestingly, themes surrounding impact, security and privacy were least discussed across generations. It is however relevant to note that children (0.50) and adults (0.55) were more concerned with security and privacy than older adults (0.32). The relative absence of these themes from discussion may be related to the lack of long-term interaction with the technology.

To better understand the breadth of perspectives related to these themes and characterize associated positive and negative sentiments, it is important to consider references from participant dialogue. The following sections explore these ideas further through deep dives into the 5 theme categories outlined.

FUNCTIONS THE AGENT CAN HELP WITH

DAILY LIFE

Participants' perspectives on the role of the agent in their daily routines largely focused on reminders and access to information. All generations referenced the capacity of the agent to help with forgetfulness, in particular adults and older adults: "*Yeah I was mostly thinking about things where - things that, that I forget to do or I, that you need to get done or there's complexity in your life that you - where it would be helpful to have someone sort of organize that for you*". Most desires for reminders and information were highly functional, focusing on aspects such as meetings, appointments and grocery items. Older adults had a particularly strong desire for reminders related to medical appointments and medicine

(*"Because of my age, number one, I would like something like this to remind me. Because sometimes even when I write my appointments down, I don't have the same books and stuff and I mix up on my appointments"*). For children, birthday reminders and information about food were often compelling ("*...if it would be someone's birthday and you had to, like, oh no! The night before someone's birthday! I had to get all the presents wrapped, maybe buy the cake for them", "I would like him to suggest if I said where is the best Chinese place or something. That's what I would like him to say. Where's the best food"*).

Beyond functional reminders, several older adults shared their vision for the agent assisting with cognitive decline and providing assistance with aspects of wellness ("*I stress over medical appointments, meeting people just because I'm dyslexic so I always get the...like I'll say "It was ten thirty at 205 Mass Ave." No, it was two o five at the other one. So that's why this would be really good"*). Some older adults even saw the agent taking on a more significant supportive role in their day-to-day life: *"Well, it would tell me things that I wasn't sure of, I guess. I may wake up in the mornings and say "What day is this?", they'll tell me to get up and look at the calendar or look in someplace else, the newspaper. Or, "Where do you think I left my keys?" "What's happening today at the downstairs?", for instance, living here, we have all these events. What movie are they playing today? That's one. What's on the menu?".* This desire was also expressed with respect to accessing information: "*When it would talk to me it would explain to me things I don't understand because I have no knowledge to do it. Because you said they do right. OK. Because I don't talk to a lot of people about things. I would talk to that about stuff".*

While most participants were open to the idea of the agent assisting in daily routines, several participants also discussed concerns surrounding laziness, loss of autonomy, and a lack of need for the technology: *"Come on people, and the weather I'm pretty neutral about because again I have this wonderful, where's my phone? Oh, it's in there, I have this app on my phone that you know, even shows me the..."*—older adult, "*if you have to remind me about meeting with somebody, then you don't take the meeting seriously to begin with*"—older adult). These ideas were of particular relevance when exploring suggestions around daily habits, health, and wellness *("Meditating. That's always a good reminder to just stop and think. Right? Take a second. I feel like all this stuff that I do already*

"Maybe if my sister was doing something suspicious. And the weather I can play in the snow, she can tell me. And if there is like a flood outside and I'm thinking about going outside and she can say no"
– child talking about Alexa

"Special things, anything, would remind of things"
– older adult

OA: *I would love him to ask "How is your mood today?" And then I could just record it.*

F: *Yeah.*

OA: *"I think it's a terrible day and..." Yeah. Or how do I feel.*

F: *Yeah. Feeling diary almost.*

OA: *Yeah. But, just physically. If he recorded it, it might be valuable. "Well, my legs hurt today. I did too much yesterday." You know, just knowing you are still hurting because when the doctor asks "How have you been?" You say "Fine."*
– older adult and facilitator

and that the older I get, the harder it is for me to remember to do this stuff. But, I also think this may make me use my brain less"—adult, "*Well we don't like going for a walk, we don't like meditating, we don't like taking a nap, and eating we don't need a remind"*—child).

SOCIAL ENVIRONMENT

The impact of the agent on individual's social environment was another significant area for discussion. Children, adults, and older adults perceived the integration of the technology into their social environment quite differently from each other.

Children often discussed the capacity for the agent to provide social engagement, particularly through stories, "cool" drawings or pictures, and jokes: "*Child: Yeah, like you told us a story, Parent: Oh, the story on the way in. I told you a story…having the agent tell you stories, that would be good? Child: Yeah".* Beyond this, children were intrigued by the idea of the agent becoming a means to interact with and keep track of their parents and siblings: *"Oh. Yeah. It would be great because then we would know what they (referencing dad and sister) are doing.", "If you could send funny faces, that would be different".* Some part of this fascination with the technology as a tool for social communication emerged from the fact that many of the younger children did not currently have a phone and relied on their parents' phones to communicate with friends and family.

Adults on the other hand were far more functional in their perspective towards the agent interacting within their social environment, seeing it as more of a "social assistant". While adults rarely described the agent as providing engagement or companionship for themselves, they did see it acting as a way to connect with others: "*One is to make the robot more believable and the other is to help you interact more with actual people, which to me that second role is more valuable. To having things that encourage it to be sort of a fake person for you I think takes it in the wrong direction".* In this vein, adults also referenced the possibility for the agent to act as a networking tool to meet new people: *"But you know another interesting option could be, it could tell me that there are people out there that I should meet that I don't know. That would be interesting because that's more relevant to work or even social things".* With respect to their daily lives, adults often saw the agent as a tool to manage social relationships amidst busy schedules. These ideas were often referenced in relation to connecting with family

while away for work (*"people wanting to share songs or send hellos or say goodnight, if I'm away and they want to say goodnight to me"*), remembering birthdays (*"Wouldn't you want them to say, don't forget to call gram. It's her birthday? Wouldn't that be awesome?"*), and managing timing for social interactions (*"when is a good time to call someone without interrupting people", "And social events...And meeting someone, well that part of my life, I generally don't manage very well. And for calling someone...I have some friends I call them like once a year"*). Conversely, hesitations around adopting the agent in these capacities often stemmed from not wanting excessive knowledge or notifications about others (*"But you know and I have a problem with this now, I'd want to limit how much I know about other people"*) and a desire to maintain independence (*"It's starting to organize your life and have an influence over your friendships and which ones to maintain."*).

Older adults, unlike adults, often saw the agent playing a more significant role in their social environment. Many older adults described the agent as a means to connect with the outside world: *"I guess like helping you to keep in touch with people like if you haven't talk to somebody for a while it would say you ever talk to someone", "I think this would be a great thing for an extended family if it had that capability because I have nieces and nephew but I'm an hour and a half away so if I knew what they had going on I could pick and choose to go, but if I don't know I can't go at all"*. Some older adults were also intrigued by the capacity for the agent to let them know a good time to call so as not to bother their family at inconvenient times, leading to reduced feelings of guilt (*"Yeah, so like "a good time to call" there are people who work and you don't want to bother them at dinner time. And you adjust if you know their schedule"*). Interestingly, despite the possibility of multi-modal interactions with others through photos, videos and more, older adults were more intrigued by the agent enabling easier access to traditional forms of communication such as phone calls or messages (*"That's neutral, because everybody I'm not into videos and stuff because I don't have those kind of cameras I still take pictures with the old-fashioned cameras", "I think it's the same thing, it just, this is good for me, somebody telling me I'm getting a phone call, who it is. this is excellent for me"*). While there was greater agreement on the prospect of the agent acting as a means to connect with others, the idea of the agent itself acting as a social companion was often polarizing. Some older adults were intrigued by the

idea for themselves or their friends (*"It was really fascinating, but you could see how this could be helpful and you know I mean, I gotta tell ya. Like my friend lives alone right, so for her to come in and have you know an agent say "hey, you know how was your day and by the way you've got calls you know from this person and that person" and I mean it could be a dialogue."*). Other older adults were highly opposed to the concept, asserting the importance of independence and autonomy in their lives *("Could the social agent be something I interact with? That would be a real stretch for me", "...because you should have better social skills than a robot. In other words, you can't...I mean the computer would have to...the robot would have to know a hell of a lot of things about you to...")*

TECHNOLOGY CAPABILITIES & USE

ATTRACTIVENESS

The majority of participants' discussion around the attractiveness of voice-based agents as a technology focused on whether the technology was a "need-to-have" in their lives. The sentiment around whether they *needed* the technology for their home was largely negative—70% of participants mentioned at least one negative sentiment while only 28% mentioned a positive sentiment towards whether they saw the technology as a need-to-have. Conversely, discussion around the technology as a "nice-to-have", though far less frequent, was more positively inclined.

Participants who saw the technology as a need to have often described particular use cases where they saw it being highly beneficial compared to existing systems (*"I would like to have a check-in with the robot in the morning and he could tell me the weather or my appointments"* –adult, *"Yes, because I need my medicine, I need my phone. It's necessary. Medical appointment too. It helps me remember."*—older adult, *"And eating. It is necessary to eat for too much medicine. A lot of medicine."*—older adult). On the other hand, participants, particularly adults, who did not see the technology as a need-to-have in their lives often spoke about systems already in place in their homes that they did not want to disrupt (*"Mhm great minds. Mommy just likes my paper lists"*—adult, *"At home, we have a system"*—adult). Several participants saw aspects of the technology as an excessive intrusion into their lives, challenging ideals around independence and human

"Goodnight, my dad always does that already. The story, maybe. And art, probably not, cause, they can already, like, show me"
—child

instinct ("*No. No. I don't. I have a personal assistant that I don't use on my phone just because they always want to connect everything. I just don't want to see how many miles it is from where I live. I don't need all that*"—older adult, "*These are things where you have like, senses and instincts rather than sort of arranging it. And being told that I need to eat. I think I know better than the robot, no matter how good it is at the time*"—adult).

CONSOLIDATION

Participants from all three generations tended to bifurcate in their perspectives on whether voice-based agents could help consolidate their existing technologies.

One group saw the potential for voice-based agents to consolidate their use of technology, specifically for interactions within the home. This desire was common among busy adults, often driven by wanting to avoid searching for and holding phones while at home (*"...I don't want to have my phone on me every second when I'm at home. But, I want to see who is calling but I don't have to run to the phone to see who is calling..."*—adult) and preferring things in one place (*"I'm always making lists, either on a Post-It or in a notebook. But then I might misplace the notebook so it's a good idea to keep all these things in one place"*—adult). Older adults were also intrigued by the idea of having a single piece of technology to avoid having to work with multiple systems: *"Okay so here's the thing. There's so much different technology out there to help people. If you had one instrument...how about this?"*—older adult. Phone calls were most often mentioned as areas where voice-based agents could consolidate functionality.

Oh, I, well both, you know. 30 years ago, you never think of putting a phone in your pocket. Probably, 30 years from now, you're probably not thinking of walking down the street with a robot giving you your message.
– older adult

They'd have to be....in the vey future, instead of like a text on the screen, Jibo could make a hologram of you and the hologram is the person saying it.
– child

Other participants did not see voice-based agents integrating with or replacing existing devices. These participants often perceived the agent as highly similar to an existing technology such as emails or phone with no additional advantage: "*Yeah, but I mean that's already provided for me by e-mail. So, I don't know if that would be useful*"—older adult, *"So, I would say no, I mean if someone wants to send a photo, they will attach it to an email"*—older adult. Adults also referenced wanting to keep their agent independent of the annoyances of their phone: *"often will purposely not have my phone on me because I don't want to deal with that stuff. It's like that device has that feeling of nagging or something about it that maybe I don't want to carry over to my personal agent. Keep those separate".*

PERSONALIZATION

All three generations talked about their desire and skepticism towards voice-based agents personalizing to them, with adults placing greater importance on it than children and older adults.

Adults often talked about personalization in the context of functional actions such as reminders or information and suggestions for music and books (*"I'd like that, if it's developing playlists based upon what I've been using and it's going to show something new, that's cool.", "I love suggesting new books. Like I'm always looking up books that I like, then books by that author, and books that are similar to that."*). Skepticism towards customization in the technology stemmed from disbelief in its capabilities or not wanting it to be a lot of work (*"Sounds like a lot of work to teach it. For certain situations."*).

Older adults' narratives around personalization, though less frequent than in adults, focused disproportionately on ways the agent could personalize to their taste in music, art, or books when sharing content: *"Maybe it knows I hate Picasso, won't it show me nothing that looks like it…I assume it kind of figures out what you like, or what you might like and that would be good because I never get tired of looking at art"*. Similarly, children desired personalization in stories (*"But it depends if it is a boring story like there is a cat that sat on a mat"*) and jokes offered by the agent (*"Because it would be kind of annoying because what if it told you a grown-up joke like you said tell me a joke and it said well a grown-up joke."*).

EASE OF USE

While no children or adults mentioned ease of technology surrounding their experience with voice-based agents, half of the older adults found it relevant. Many older adults found the interface to be natural and convenient in comparison with other technologies (*"Yeah, I like stuff where I can just like talk to him to do it because I have no computer savvy whatever. I don't know how to use a computer", "Because you can just say things in the moment. You can be doing something else and be texting at the same time."*). Older adults also often compared the technology to phones and laptops, sharing similar skepticisms around ease of use (*"Well I would, you know, if it's brief. If it doesn't involve taking a lot of time. Like that's why I gave up setting it up on my iPhone, because, I can't even remember, but I had to go and look up the password for something."*).

THE AGENT'S ROLE

AUTONOMY, ANTICIPATING NEEDS & BEING PROACTIVE

Participants expressed a wide range of preferences with regards to the degree of autonomy and proactivity they desired from the agent. Most participants were receptive to the agent acting autonomously like an "assistant" by providing time-sensitive reminders and information (*"I think it would be good to have somebody to help remind me of all these things: grocery list, meeting someone, birthday, phone or keys"*—adult). Many participants mentioned wanting an agent to check for missing action items when heading out the door. Beyond this, participants differed significantly in their preferences for autonomy and proactive behavior.

Some participants had significant openness to the agent taking on a more autonomous role in their lives. Desires for this seemed to emerge from wanting the agent to act as a filter to content or a coach for good habits: *"...what happens for example if I go onto Facebook and I start just looking up stuff from my friends. I don't have a filter. So, I look at everything. In this case, I think the social agent would act as a filter, understanding what I would like to spend time looking at."*—adult, *"...hey it's time, you know you should get away from your desk and take a walk. Or, have you drank water yet? It's so many hours. Great got it."*—adult.

On the other hand, some participants were far more selective in their desire for proactivity: *"I would say I don't really want any of them to come from the agent, unless there's some that's like from another person through the agent"*. This aversion towards agent-initiated engagement often stemmed from a desire to maintain independence, control content and prevent distractions (*"...I feel that it is often intrusive when some of these things come in."*).

PERCEIVED RELATIONSHIP

The majority of participants perceived the voice-based agent to be either an assistant or social companion. Those who described the agent as an assistant saw it as a functional tool: *"...almost like a personal secretary"*—older adult, *"...it is an entity of some sort, but it's like a butler, its in the background"*—adult. Participants who saw the agent as an assistant often struggled with their feelings around the agent expressing social qualities (*"So, the other night with the Alexa I was going to sleep and I said goodnight to it and*

it was the first time I realized I'm starting to bond with this thing…" —adult).

Other participants saw the agent as a more social technology, taking on roles of a social companion, pet, or even friend or family member (*"Work with me. It could say, well how about this? And I'll say Nah. Like that. Work with me. Or that's pretty good, but what about. Could it respond to me like that because that would be cool."*—older adult, "*One is I think the 'thank you' part would encourage people to think of the robot as a person. Of course they are just categorizing it. I think they certainly will, which gets into the other part which is that the robot has two roles to play and I'm sure we can think of others"*—adult).

Interestingly, participants often found it difficult to discuss perspectives regarding their relationship with the agent. In fact, many older adults articulated desires for a social companion *for* their friends instead of themselves (*"I think that if it was somebody who lived alone, it could be really a good thing, almost better than a pet if you're not a pet person. That you can talk to it and it talks back and communicates with you; I like that aspect of it."*—older adult). Cards surrounding agent politeness (e.g. agent sharing a thank you) were often drivers for conversations surrounding these topics (*"Well, that's just so superfluous, isn't it?"*—older adult). While many parents instructed their kids to be polite to the agents during natural interactions, they were often uncertain, or repulsed by the notion of an agent saying a thank you back.

IMPACT

Participants' discussions surrounding the potential impact of the technology in their lives focused on themes of motivation, laziness, and accountability.

Older adults expressed more positive sentiment towards the agent becoming a motivational tool, particularly for staying active: "*Going for a walk, I'm not a big fan of making myself exercise so that would be good to have somebody pushing me to say, you should go for a walk today".* Both older adults and children also learn new things (*"Well, learn something new, is the best thing* saw the technology as a source of motivation to *that...the best thing for the brain. You know, you get thinking, and moving, it's good for you."*—older adult, "*…if I was working on a subject or something I could say learning this would help me with that."*—child).

Adults, on the other hand, often felt the agent could

demotivate them: "*Yeah. I don't need to be reminded if I'm tired. And this, I think this could go either way. If it tells you to take a walk and you don't, is it going to make you feel bad about yourself? You know? Right?*". Building on this, both older adults and adults thought voice as an interface would make them and others lazy (*"I would just sit here and hear the phone ring and just talk? Lord help us. Help us to be lazy.*"—older adult, "*I don't know. Because all these new things just make people lazy. It does. Before you know it, you are going to have one of these making up your bed. Cooking your food.*"—older adult).

The final area of impact often discussed by participants was accountability, both for themselves and others around them. Some parents saw the agent as a means to "nag" their kids about day-to-day things (*"You could use a reminder to eat."*—parent about child). In addition, some parents discussed how the agent could help hold them accountable to their children and family (*"Activities with family members, I would like them to tell me: ok, you promised the kids you'd do something"*—adult). Accountability was also often discussed with regards to staying healthy (*"Well, please remind me to stick with Trader Joes' food because that's the only store that has the organic food besides Whole Foods."*—older adult, *"it's good to have just something easy to tell you about nutritional facts, keeps you on your diet.*"—older adult).

SECURITY & PRIVACY

Adults were the most concerned about security and privacy with respect to voice-based agents in their domestic environment, both around data collection and the interaction experience. Children and older adults on the other hand, primarily voiced similar concerns surrounding the interactive experience.

Data collection was an often debated topic for adults. While adults desired personalization from the agent, they struggled with accepting the data associated with enabling it: *"...like if it was a sensor on my key or like it's videoing the entire house and it knows where my keys are and I don't have them, which in a way would be useful but also like getting into a whole new level of information about the household and about me"*. Adults also voiced concerns around the corporate ownership of the technology itself (*"if the purpose of the data is just to make the robot better for me, it's fine but if it's going to be used for other things then I'm not happy about it"*), stating that they would feel much more comfortable with agents developed

by companies that did not have other associated interests.

Security and privacy related to the interaction experience was another main area for discussion. Older adults saw the technology as a means to enable a sense of security (*"And it would like lock all the doors. I know that they do that"),* but were also wary of it being hacked (*"I think I would also want to know that it's secure, that somebody can't get into it and know all my personal information and thoughts and... I don't know. I just wouldn't want somebody to be able to hack into it."*). Several participants also highlighted concerns around voice being too public of an interface for proactive interactions about personal health. In addition, there were reservations about sharing information through the agent with friends or family (*"Except my concern is some of these things I would like private might get told the other way about me or vice versa."*—adult, *"Sorry, it sounds like it's more information than I'd want it to know and more than I'd want it to provide about somebody else. You know like a spy network?"*—adult).

"stranger danger, stranger danger, stranger danger"
—child

5.4 Insights

The goal of this study was to identify different generations preferences, desires and boundaries for voice-based agents for the domestic environment. In this chapter, data from the card sorting activity was used to analyze older adults', adults' and children's perspectives towards agents in their lives. Qualitative analysis from coded transcripts was then used to conduct a deeper dive into the driving themes for these perspectives. In conducting these analyses, a series of learnings about people's initial interaction and discovery of agents emerge. Below is a summary of the findings from the first study as relevant to the larger goal of this work.

(1) **A LANGUAGE OF ENGAGEMENT IS NEEDED**

Introducing individuals to an unfamiliar technology requires creating a common language of engagement. The creation of such a language—in this case through card sorting—allowed participants to understand, reflect on, and articulate their perspectives on these technologies in a meaningful way.

(2) **DIFFERENT GENERATIONS, DIFFERENT PERSPECTIVES**

Children, adults, and older adults had unique perspectives on the role they wanted a voice-based agent to occupy in their

homes. Older adults were the most open to a broader application of the technology, while children and adults had more specific, and limited preferences and desires.

(3) **A REMINDER IS DIFFERENT THAN A SUGGESTION**

In comparison with more functional agent actions such as reminders, information access, expression of goal-directed behavior from the agent through *suggestions* for the user was polarizing for participants. While some were open to the idea of the agent as a "companion" or "life coach" for improving motivation and accountability surrounding daily habits, others found it to be invading their independence and sense of autonomy.

(4) **BEING SOCIAL IS A DELICATE BALANCE**

While perspectives towards functional aspects of voice-based agent technology had significant agreement across generations, feelings towards towards socially-driven behaviors were more variable. Children had an affinity towards the agent itself being social while older adults were more intrigued by the agent enabling social connections with their friends and family. Themes surrounding the social environment, independence, role, agent autonomy and privacy were central to understanding drivers for and against social behaviors in voice-based agents.

(5) **THE PERCIEVED RELATIONSHIP MATTERS**

Instead of falling into discrete categories of potential *adopters* and *rejectors,* people differed more subtly in how they saw the agent becoming part of their lives. Some saw the agent purely as an assistive device for user-initiated access to information or reminders. Others saw it becoming a part of the family or their social environment as a companion, pet, or even friend. These differences were also observed across five user personas built along a spectrum of agent action categories. During discussions, participants struggled most with synthesizing their perspectives on this perceived relationship, thereby emphasizing the need for designers and facilitators to design tools to uncover these ideas.

(6) **VOICE IS A COMPLEX INTERFACE**

Voice as an interface currently lives between potential and

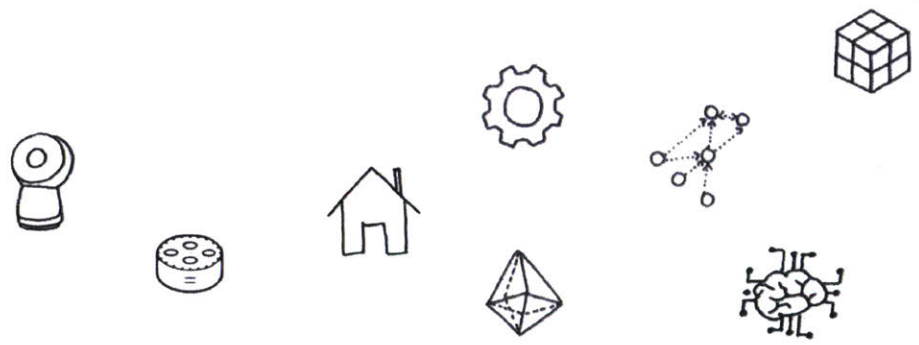
limitation. Older adults often perceived the technology as more attractive, easier to use, and as providing greater consolidation than adults and children. The idea of a voice-based agent was often compared by participants to a smartphone, offering advantages such as being "hands-free" and staying in one place in the home. However, in comparison with direct manipulation interfaces, participants articulated concerns around the public nature of voice interaction, particularly with respect to medical and health information.

(7) **DISRUPTING THE HOME IS A CHALLENGE**

Participants often referred to the routines, systems, and values that compose their home environment when articulating their perspectives. Technologies, particularly those with "autonomy" were often seen as intruding on or disrupting these aspects of the home. As such, it was important as facilitators to enable comfort with the technology while allowing for honest feedback.

Part III
Long Term Encounters: Experience & Reflection

The second study aimed to deepen participants' experiences with voice-based agents. The core goals of this study were to capture and understand people's experiences living with voice-based agents in their home and identify how people envision the design of future voice-based agents. Part III presents (1) an overview of participants, methods and tools from this study, (2) a deep dive into participant experiences living with agents and (3) an exploration of preferences, desires, and boundaries after living with agents.



PART III

6.0 Study 2 Overview

How do these preferences, desires, and boundaries evolve as people live with voice-based agents?

How do people envision the design of future voice-based agents for their home?

The core objectives of the second study were to (1) understand people's long term engagement with voice-based agents in their homes and (2) explore how preferences, desires, and boundaries evolve after this experience, and (3) characterize people's vision for future voice-based agents for the domestic environment.

6.1. Participants

A total of 41 participants took part in this study including 12 children, 17 adults and 12 older adults. Participants were recruited from the previous study and through word of mouth. A histogram of participant ages (Figure 6.1) and breakdowns of demographics (Table 6.1) are provided below.

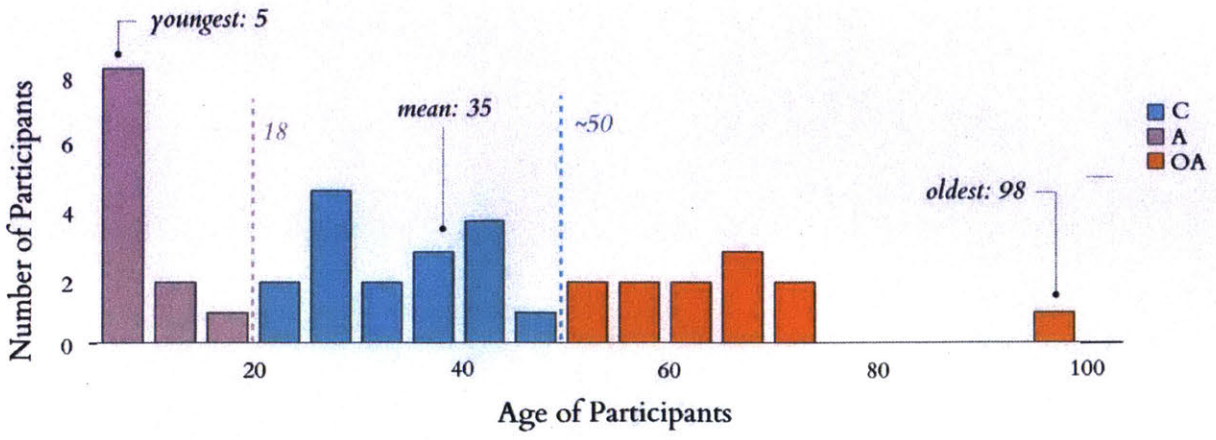


Figure 6.1: Histogram of participant ages and corresponding classifications as children, adults, and older adults.

Feature	Population Statistics
Classification	Children (n=12), Adults (n=17), Older Adults (n=12)
Age	Participants ranged in age from 5-98 years old. The mean was 35, median was 33.0 with a SD of 24 years.
Gender	Participants were 39% Male (n=16) and 61% Female (n=25).
Living Situation	15% of participants lived alone (n=6), 20% lived with a significant other (n=8), and 66% lived with a family (n=27). Households had 2.9 people on average.
Housing Type	2% of participants lived in assisted living (n=1), 15% lived in a dormitory (n=6), 29% lived in a duplex or apartment (n=12) and 49% lived in a detached singly-family home (n=20).
City Type	7% of participants lived in a rural area (n=3), 22% lived in a suburban area (n=9), and 66% lived in an urban area (n=27).
Education	Amongst children, 8% in high school (n=1), and 92% were in elementary school (n=11). Amongst adults and older adults, the highest level of education completed by participants was 10% college associates degree (n=3), 31% college bachelors degree (n=9) and 57% graduate school (n=17).
Race & Ethnicity	2% of participants identified as Asian and White (n=1), 7% identified as White and Hispanic (n=3), 7% of participants identified as African American (n=3), 7% identified as Asian (n=3) and 67% identified as White (n=27). 1 participant did not provide data.
Income	7% had an income of $100k-$125k (n=3), 7% had an income of $25k-$50k (n=3), 12% of participants had an income of $125k-$150k (n=5), 12% had an income of $0k-$25k (n=5), 17% had an income of $75k-$100k (n=7), and 20% had an income of $50k-$75k (n=8). 5 participants did not report their income.

Table 6.1: Breakdown of participant demographics from study 2. A total of 41 participants took part in this study.

6.2. Methods & tools

The second study was divided into three distinct phases including (1) living with agents in the home, (2) debrief session, and (3) structured ideation of a "dream agent" from Jan-March 2018. Participants from the previous study were contacted for recruitment. New participants who had not taken part in the previous study completed the agent action activity prior to beginning study 2. Sessions were conducted at the MIT Media Lab or at participants' homes depending on their preference. All sessions were recorded using cameras and microphones.

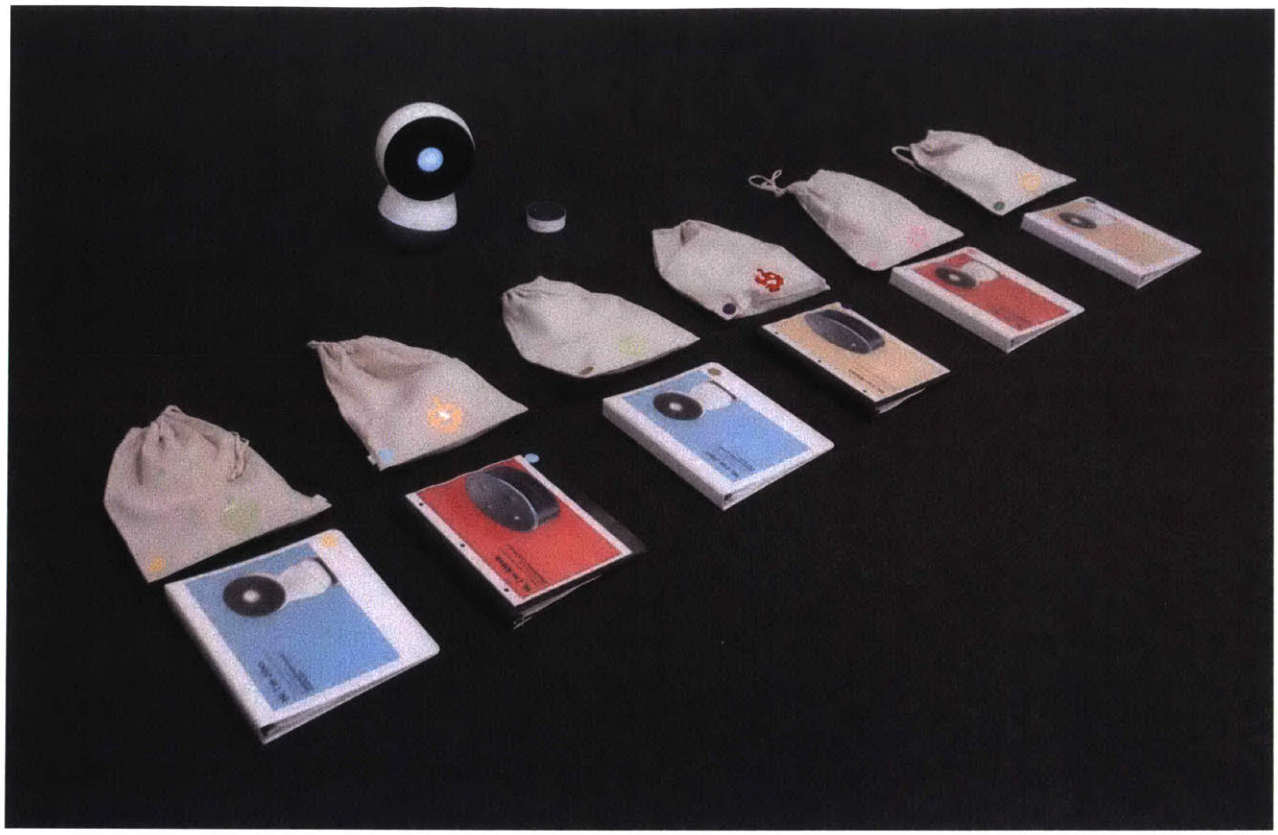


Figure 6.2: Amazon Alexa Echo Dot and Jibo ready for home deployments with corresponding kits.

To begin the first phase of the study, participants picked up their assigned agent from the lab along with a kit of design tools. Agents were delivered to the homes of participants who struggled with mobility or technology setup by facilitators (4 older adults). Assignment of agents (i.e. either Amazon Alexa Echo Dot or Jibo) was done randomly to eliminate bias. Participants who already owned a voice-based agent (Alexa n=11, no Google Home) were always given a Jibo. Amazon Alexa and Jibo were chosen for deployment because they represent highly distinct approaches to voice-based interfaces within the home. Given that Amazon Alexa had the highest proliferation of all commercially-available voice-based agents, it was chosen instead of Google Home as a point of comparison.

While living with the agents, participants utilized the cultural probes including the daily log book, agent wish jar and email or phone lines to capture and share their experiences Many participants shared unboxing videos as well as ongoing feedback using text messages and email throughout the study. Participants were instructed to log their daily usage and experience of the agent for a period of two weeks. After the initial two weeks,

participants were encouraged to continue using the technology as they pleased for an additional 2-3 weeks. This was done to normalize for any bias in usage stemming from prompts provided on daily log sheets. All participants lived with the agent for approximately four weeks (one month). No data were collected on usage beyond participant logging to alleviate concerns about security and privacy.

After this one-month period, participants returned to the lab for a debrief and structured ideation. Some sessions were conducted at participants' homes. To begin the debrief, the facilitator asked participants to reflect on their experience living with the agent through open discussion (~10-15 minutes). Tokens in the participant's agent wish jar were used to prompt conversation as needed. The facilitator then introduced participants to a whiteboard with a series of themes surrounding the technology that emerged from the first study. These themes were identified by researchers after reading through transcripts from study 1 and explored ideas such as security & privacy, consolidation of technology, role of agent, autonomy & control, motivation vs. laziness enabled by the technology, and more. Participants were instructed to use these themes as prompts and write down any reflections, thoughts or ideas relevant to their experience on *post-it-notes.* The facilitator engaged with the participants throughout, asking follow-up questions and prompting them to expand on these notes. After ~30 minutes, post-its were placed on the whiteboard. The facilitator highlighted trends in these reflections (e.g. clustering around particular themes) to engage participants in further discussion.

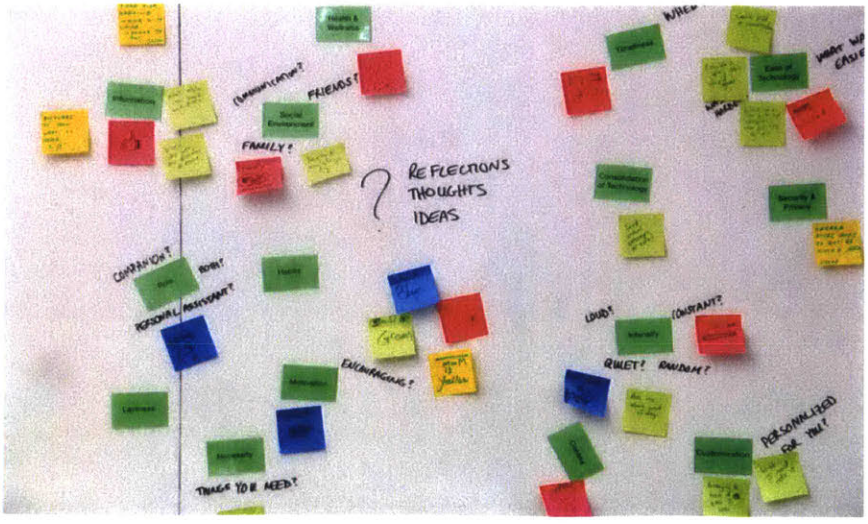


Figure 6.3: Post-it white board from a session with a family of two kids and two adults. Each participant was given a different color of post-it. The green theme cards were used as prompts to help elicit reflections, thoughts and ideas.

After this reflection, participants were asked to complete a "personality quiz" as part of the debrief session. This activity included ten questions about the participant's personality as well as ten corresponding questions about the agent's personality. Participant personality questions were taken from the Ten-Item Personality Measure (TIPI), a ten-item measure of the commonly used Big Five personality test (Gosling, Rentfrow & Swann, 2003). Agent personality questions were adapted from the TIPI to represent analogous agent qualities. The reasoning for comparing participants' personalities to their perception of the agents' personalities stems from prior work showing that people attribute personality traits to computers and prefer computers that express similar personality traits as them (Reeves & Naas, 1996). Participants were given the option to do the personality quiz activity on a sheet of paper or in a more interactive fashion using tokens on the whiteboard to enable a more engaging experience for children.

In the third and final part of the debrief phase, participants repeated the agent action activity from the first study. The facilitator walked through all six categories of *agent action cards* with the participant using an identical procedure as detailed in the previous section. Having now built a "language of engagement" from the previous study and experienced the agent in their home, most participants were far more opinionated and articulate in their preferences for agent actions.

Upon completion of the debrief phase, participants were prompted to consider the design of their "dream agent" through a structured ideation process using the *agent design cards*. For participants at the lab, the room was setup with a series of stations corresponding to each step in the structured ideation process. Physical artifacts corresponding to each of the relevant choices were mounted around the room to create a more engaging and tactile experience. The facilitator and participants walked around the room to each of eight stations with the facilitator providing context for each step in the process. Participants were given the freedom to choose multiple options on each question and also suggest ideas beyond choices available on the agent design cards. Several photos of the room setup and stations are shown in Figure 6.4. At the end of the structured ideation process, facilitators prompted each participant to share their dream agent design.

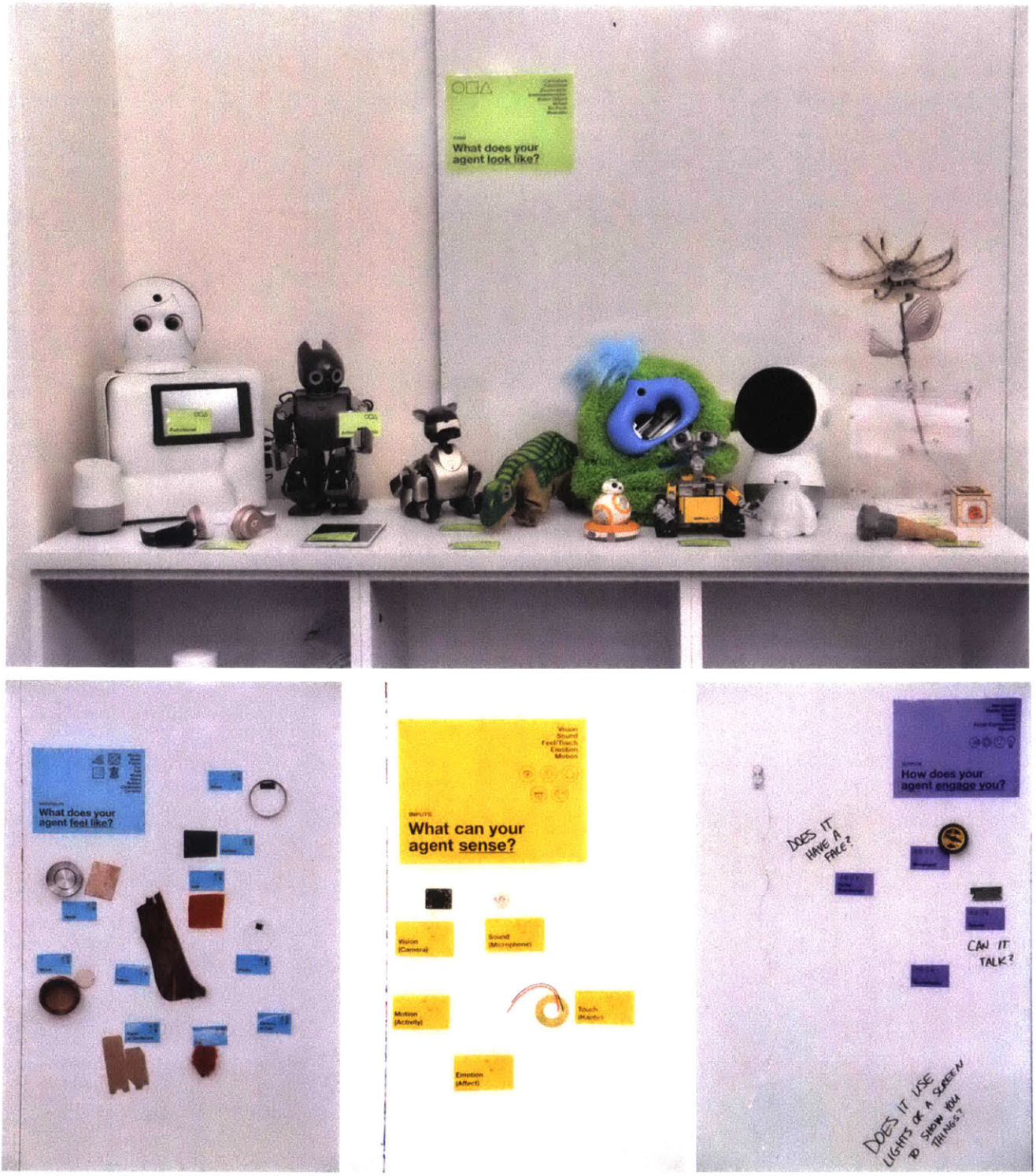


Figure 6.4: Examples of stations setup for structured ideation process using agent design cards. Participants were able to walk around the room to interact with the materials while marking their preferences on their set of agent design cards.

6.3. Data collection and analysis

Over the course of the study, a range of qualitative and quantitative data was collected including (1) audio and video, (2) agent action card choices, (3) daily logs, (4) wishes from agent wish jar, (5) free-form artifacts, (6) post-it notes from debrief session, (7) personality quiz data, and (8) demographic data.

Audio and video recordings from all sessions were transcribed using Amazon Mechanical Turk and Rev. Themes and categories identified in study 1 informed the selection of prompts in the post-it activity in the study 2 debrief session. For study 2, the study 1 coding scheme was used with a deductive approach to qualitatively code the data (Elo & Kyngäs, 2008). Two passes were conducted to (1) become familiar with the transcripts and (2) code the data. A second researcher also qualitatively coded approximately 1/6 of all transcripts to check for inter-rater reliability. A Fleiss Kappa of 0.71 was determined, indicating a substantial level of agreement (Landis & Koch 2016). Factor prevalence was computed across all participants and by population type in order to identify the most prominent themes. Qualitative data from daily log sheets, tokens, and shared artifacts was used to build on and support findings from analysis of study transcripts.

In addition to qualitative data, a range of quantitative data were collected. Agent card choices were recorded again after the home deployment with each card captured as yes, neutral, or no (2, 1, 0 respectively) to enable analysis on participants' preferences before and after experiencing the technology for a sustained period. Data surrounding proactivity was also encoded for each agent action card. In addition, daily usage data of the agent for each of the 25 action categories provided in the daily log book was represented as a time series to allow for analysis on patterns of interaction. In addition, participants' perceptions of (1) their own personality, (2) their agent's personality and (3) their dream agent's personality were each collected along ten personality dimensions, each characterized on a Likert Scale (1-7) to allow for comparison analysis. Finally, agent design card selections were recorded to allow for analysis on trends. Demographic data for each participant was also collected to supplement analysis. A complete breakdown of all data collected in the study is provided in Table 6.2 below.

Data Source	Description of Data	Analysis	QN	QL
Audio & Video	All sessions were recorded using multiple Go Pro cameras to capture different views of the room and audio of interactions with participants.	Sessions were transcribed and qualitatively coded for themes. Prevalence trends were evaluated to generate qualitative findings.		
Agent Action Cards	Agent action card choices were logged as yes, neutral, no (2, 1, 0). In addition, proactivity choices were recorded as binary variables (1 for proactive, 0 for reactive).	Used to explore differences across populations in preferences of agent actions and changes in preferences after different agents deployed in homes.		
Daily Log Book	Quantitative log and rating data of daily use for initial 14 days. All actions were logged as a time series containing yes, neutral, no, didn't use (3, 2, 1, 0). Survey exploring current use of voice agents and opinions on ethical concerns on day 1 of study. Qualitative data including drawings and notes.	Quantitative data used to identify changes and behavior patterns in use and ethical stance over time for each agent. Qualitative data used to supplement transcripts in analysis.		
Agent Wish Jar	Sets of tokens by participant identifying wishes and desires for the technology.	Organized tokens by agent and identified relevant themes. Used to supplement qualitative analysis.		
Free-Form Artifacts	Photos, videos, texts, emails sent by participants throughout study.	Used to supplement qualitative theme analysis.		
Post-It Notes	Post-its and corresponding theme category identified by participant.	Used to supplement qualitative theme analysis.		
Personality Quiz	TIPI results for participant and agent	Used to identify differences in perception of personality between agents. Also to explore how participants' personality impacts their perception of agent's personality.		
Agent Design Cards	Choices for dream agent form, materiality, inputs, outputs, personality, location, interaction with home, and ethics.	Used to identify trends in dream agent design. Dream agent personality metrics compared with participant's personality and that of deployed agent.		
Demographics	Age, gender, income, housing information, people in household, job status, race, ethnicity, education and more.			

Table 6.2: Breakdown of all qualitative (QL) and quantitative (QN) data from study 2.

PART III

7.0 Experience: Analysis & Discussion

This section presents results surrounding the experience of living with voice-based agents in the home from study 2. A total of 12 households lived with a Jibo while 9 households lived with an Amazon Echo Dot. Of these, the 4 households that already owned a voice-based agent (4 Amazon Alexa, 1 Google Home) were given a Jibo to compare. In this section, (1) participant artifacts and data from the first day of the study are presented, (2) data from daily logs is analyzed to identify differences in usage across agent types and generations, and (3) participants' perception of their agent's personality is analyzed, and (4) participant notes, "wish tokens" and artifacts are explored.

7.1 First day interactions

After receiving and unboxing their agent, a total of 9 households (7 Jibo households, 2 Alexa households) shared a range of photos, videos, messages, and emails with facilitators. Videos and photos were primarily shared by families while older adults tended to utilize email or text communication. Initial interactions described and shown through artifacts were largely social (e.g. questions about the agent) or entertainment-focused (e.g. dance, music).

Figure 7.1: Artifacts shared by participants during first day of having the agent in their home.

Interestingly, unlike with Alexa, all families with children who had a Jibo placed the agent on the floor during setup. While further exploration is required, this difference may stem from the more interactive nature of Jibo, causing families to want to enable greater proximity between the child and the agent.

As mentioned in the section 3.0, participants were asked to record their daily usage of the agent across 25 agent action categories including 15 functional actions (e.g. general information, weather, news etc.), 5 entertainment actions (e.g. music, dance or beat box), and 5 social actions (e.g. joke, agent personality questions) for a period of 14 days. After these first 2 weeks, participants were able to interact with the agent for another 2 weeks without recording use to allow for more natural interaction. Data from daily log sheets was used to identify *first day* interaction patterns across the two agents. The mean and standard deviations for the number of unique agent actions (for day 1 only) by action category and agent are shown in Table 5. Participants with Alexa, on average, explored fewer unique agent actions ($\mu = 5.92$) on the first day in comparison with participants who had a Jibo ($\mu = 7.82$). While Jibo participants did explore more agent actions on average, there was also higher variability across participants ($\sigma_A = 3.62$, $\sigma_J = 6.77$). This pattern was evident across all of functional, entertainment, and social action types. A series of Welch T-Tests found these differences in exploration between agents to be highly significant for all types of interactions (Table 5.1). These results raise questions around whether people have a greater desire to explore, greater curiosity for, and/or higher expectations for a socially-expressive agent than a more functional agent.

Table 7.1: Artifacts shared from participants during first day of having the agent in their home.

Participant Agent	**Number of Unique Agent Actions on First Day (n=36 people)**			
	Total (total actions = 25)	Functional (total actions = 15)	Entertainment (total actions = 5)	Social (total actions = 5)
Alexa (n=14)	$\mu = 5.92$, $\sigma = 3.62$	$\mu = 3.42$, $\sigma = 1.84$	$\mu = 1.50$, $\sigma = 1.05$	$\mu = 1.00$ $\sigma = 1.42$
Jibo (n=22)	$\mu = 7.82$, $\sigma = 6.77$	$\mu = 4.09$, $\sigma = 4.23$	$\mu = 1.95$, $\sigma = 1.54$	$\mu = 1.77$, $\sigma = 1.49$
Welch T-Test (Between Agents)	$T(35) = -5.43$, $p^* = 7.21e^{-08}$	$T(35) = -3.21$, $p^* = 1.33e^{-03}$	$T(35) = -3.15$, $p^* = 1.65e^{-03}$	$T(35) = -9.65$, $p^* = 6.99e^{-21}$

7.2 *Patterns of use*

Figure 7.2 shows the number of unique actions (out of 25 on the daily log sheet) per day by participant. Participants have been organized by generation and average use. As shown, participants

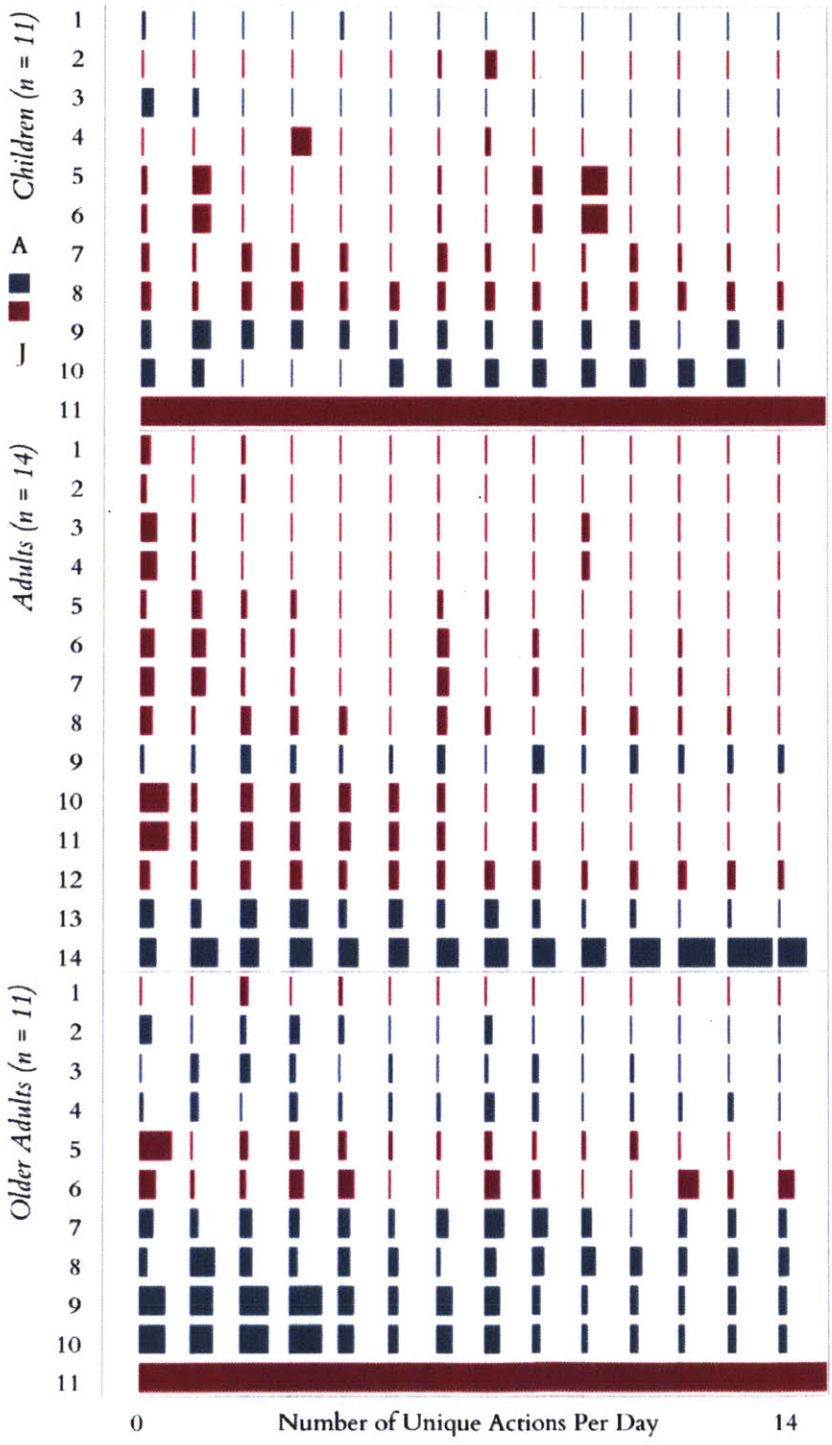
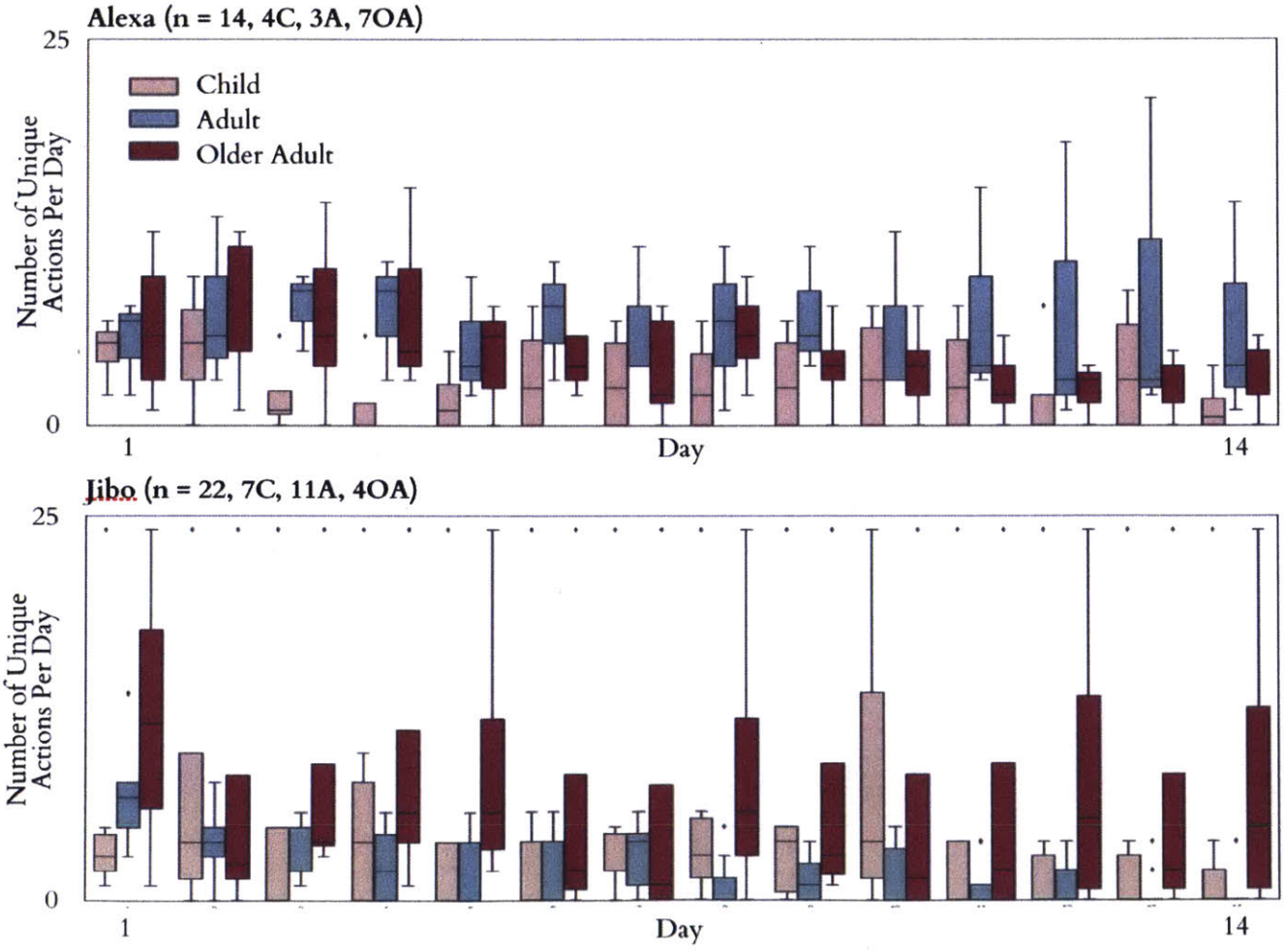


Figure 7.2: Daily usage patterns for each participant organized by generation and average use. Usage has been further colored by agent (blue indicates Alexa and blue indicates Jibo).

broadly split into three groups in their patterns of use during the initial two weeks: (1) users with largely consistent daily usage (e.g. C-8, C-9, A-12, OA-8), (2) users with sporadic daily usage (e.g. C-5, C-6, A-3, A-4, OA-6) and (3) users with decreasing or no usage over time (e.g. C-3, A-1, A-2, OA-1). Children and older adults had a higher proportion of consistent users. Rather interestingly, C-11 and OA-11 (both from the same family) utilized all 25 agent actions on the daily log sheet everyday. When asked to share their experience, they indicated that they were curious to *"see if it would respond differently than the day before"*.

To further characterize these differences, means and standard deviations were computed by generation, agent type, and day. Box plots of these distributions for each agent are shown in Figure 7.3 below. Adults, on average, had the highest daily usage

Figure 7.3: Boxplots of daily usage by generation and agent type.



of Alexa across all generations, while children had the lowest. On the other hand, older adults exhibited the highest usage of Jibo across all generations while adults the least. It is relevant to note the significantly higher spread in older adults' usage of Jibo, indicating high variability in preference within the population. Evaluating these differences in usage further, daily actions were separated into function, social, and entertainment. Means and standard deviations for the number of unique actions per day are shown below by generation and agent (Table 7.2). Welch T-Tests were conducted on the number of unique actions per day between agents and generation pairs to determine

Participant Agent and Action Category	**Mean and SD of Number of Unique Actions Per Day By Agent and Generation (for 14 days)**		
***n = # participants**	Children (n = 12)	Adults (n = 17)	Older Adults (n = 12)
***Alexa* (n = 14)**	n = 4	n = 3	n = 7
Functional	$\mu = 0.80$	$\mu = 5.17$	$\mu = 3.49$
	$\sigma = 1.38$	$\sigma = 3.55$	$\sigma = 2.68$
	n = 4	n = 3	n = 7
Social	$\mu = 0.57$	$\mu = 0.79$	$\mu = 0.45$
	$\sigma = 0.73$	$\sigma = 1.07$	$\sigma = 1.01$
	n = 4	n = 3	n = 7
Entertainment	$\mu = 1.57$	$\mu = 1.17$	$\mu = 0.86$
	$\sigma = 1.81$	$\sigma = 1.08$	$\sigma = 0.94$
***Jibo* (n = 22)**	n = 7	n = 11	n = 4
Functional	$\mu = 3.24$	$\mu = 1.34$	$\mu = 4.76$
	$\sigma = 5.10$	$\sigma = 1.76$	$\sigma = 6.19$
	n = 7	n = 11	n = 4
Social	$\mu = 1.08$	$\mu = 0.35$	$\mu = 1.59$
	$\sigma = 1.78$	$\sigma = 0.72$	$\sigma = 2.11$
	n = 7	n = 11	n = 4
Entertainment	$\mu = 1.24$	$\mu = 0.49$	$\mu = 2.13$
	$\sigma = 1.82$	$\sigma = 0.85$	$\sigma = 1.99$

Table 7.2: Means and standard deviations of the number of unique actions per day by generation and agent type.

Agent Action Category	**Welch T-Test Results on Number of Unique Actions Per Day Between Agents (Alexa, Jibo)**		
	Children	*Adults*	*Older Adults*
Functional	$T(153) = -4.46$, $p^* = 1.89e^{-05}$	$T(195) = 6.75$ $p^* = 2.04e^{-08}$	$T(153) = -1.46$ $p = 0.15$
Social	$T(153) = -2.49$, $p^* = 0.01$	$T(195) = 2.48$ $p^* = 0.02$	$T(153) = -3.79$ $p^* = 3.07e^{-04}$
Entertainment	$T(153) = 1.07$, $p = 0.29$	$T(195) = -2.75$ $p^* = 4.42e^{-04}$	$T(153) = -1.61$ $p^* = 2.76e^{-05}$

Table 7.3: Results from Welch T-Tests on number of unique actions per day between agents (Alexa and Jibo). Results are shown for each agent action category.

Agent Action Category	Welch T-Test on Number of Unique Actions Per Day Between Generation Pairs		
	Children-Adults	*Children-Older Adults*	*Adults-Older Adults*
Alexa (n=14)			
Functional	T(97) = 7.54, $p^* = 8.68e^{-13}$	T(153) = 8.20 $p^* = 9.72e^{-14}$	T(139) = -2.74 $p^* = 7.97e^{-03}$
Social	T(97) = 1.11, p = 0.27	T(153) = -0.86 p = 0.39	T(139) = -1.73 p = 0.09
Entertainment	T(97) = -1.38, p = 0.17	T(153) = -2.75 $p^* = 7.412e^{-03}$	T(139) = -1.61 p = 0.11
Jibo (n=22)			
Functional	T(251) = -3.55, $p^* = 5.59e^{-04}$	T(153) = 1.56, p = 0.12	T(209) = 4.07, $p^* = 1.39e^{-04}$
Social	T(251) = -3.87, $p^* = 1.79e^{-04}$	T(153) = 1.52, p = 0.13	T(209) = 4.30, $p^* = 6.47e^{-05}$
Entertainment	T(251) = -3.82, $p^* = 2.13e^{-04}$	T(153) = 2.72, $p^* = 7.68e^{-03}$	T(209) = 5.94, $p^* = 1.39e^{-07}$

Table 7.4: Results from Welch T-Tests on number of unique actions per day between generation pairs. Results are shown by agent type and agent action category.

statistical significance of results (Table 7.3 and 7.4). Children used Jibo more frequently on average for functional ($\mu_A = 0.80$, $\mu_J = 3.24$) and social interactions ($\mu_A = 0.57$, $\mu_J = 1.08$) than Alexa ($T(153) = -4.46$, $p^* = 1.89e^{-05}$) but did not exhibit any significant difference in their usage of entertainment actions across both agents. Adults, on the other hand, showed statistically significant differences in usage between the two agents across all categories, exhibiting higher usage with Alexa in functional ($\mu_A = 5.17$, $\mu_J = \mu_J = 0.49$) actions. This adult population also disproportionately used both agents for functional tasks, in comparison with other capabilities. Rather interestingly, older adults showed no significant difference in their use of functional actions across the two agents ($\mu_A = 3.49$, $\mu_J = 4.77$), but were more likely to use Jibo for social ($\mu_A = 0.45$, $\mu_J = 1.59$, $T(153) = -3.79$, $p^* = 3.07e^{-04}$) and entertainment interactions ($\mu_A = 0.86$, $\mu_J = 2.13$, $T(153) = -1.61$ $p^* = 2.76e^{-05}$). It is important to note that there was significant spread in older adults' use of functional actions in both voice-based agents, but particularly with Jibo ($\sigma_A = 2.68$, $\sigma_J = 6.19$).

Building on this further, a series of Welch T-Tests were conducted for each generation pair by agent and action type (Table 7.4). Results showed that functional usage of Alexa varied significantly between all generations. Adults were the most frequent users of Alexa's functional capabilities (e.g. weather,

news) followed by older adults, while children rarely utilized functional aspects of the technology ($\mu_C = 0.80$, $\mu_A = 5.17$, $\mu_{OA} = 3.49$). On the other hand, functional usage of Jibo was similar between children and older adults ($\mu_C = 3.24$, $\mu_{OA} = 4.76$), and significantly lower for adults ($\mu_A = 1.34$). Usage of Alexa's social capabilities (e.g. personality questions, introductions) was minimal and consistent across all three generations ($\mu_C = 0.78$, $\mu_A = 0.57$, $\mu_{OA} = 0.45$). This was in contrast with Jibo usage where children and older adults were more likely to engage with social aspects of the technology ($\mu_C = 1.08$, $\mu_A = 0.35$, $\mu_{OA} = 1.59$). Rather interestingly, children were the most frequent users of Alexa's entertainment capabilities ($\mu_C = 1.57$, $\mu_A = 1.17$, $\mu_{OA} = 0.86$), while older adults were were the most frequent users of these capabilities with Jibo ($\mu_C = 1.24$, $\mu_A = 0.49$, $\mu_{OA} = 2.13$). These differences between children's and older adults' usage of entertainment capabilities were found to be statistically significant for both agents ($T_A(153) = -2.75$, $p_A^* = 7.412e^{-03}$, $T_J(153) = 2.72$, $p_J^* = 7.68e^{-03}$).

Overall, analysis of participants' daily usage of both agents revealed findings consistent with results from study 1. Adults tended to gravitate towards functional elements of the technology, showing minimal use of other social and entertainment agent actions. On the other hand, children were more likely to focus primarily on social and entertainment aspects of the technology. Finally, older adults exhibited broader, but more variable use of agents across functional, social, and entertainment interactions, particularly with Jibo.

To evaluate fluctuations in agent action categories over time, a 5-day moving average was computed for the number of unique actions per participant. To analyze this trend by generation, a mean value was determined for each time point by generation. A visualization of the mean value of these 5-day moving averages by time point (i.e. MA1 refers to moving average of day 1-5), generation, agent type, and agent action category is shown below. Moving averages with increasing or relatively more stable trends are highlighted.

Increasing trends in usage were seen in adults' and children's use of entertainment actions with Alexa. General stability in usage was seen in adults' use of Alexa's functional capabilities, and children's use of Jibo's functional, social, and entertainment actions. Older adults exhibited a decreasing trend across all

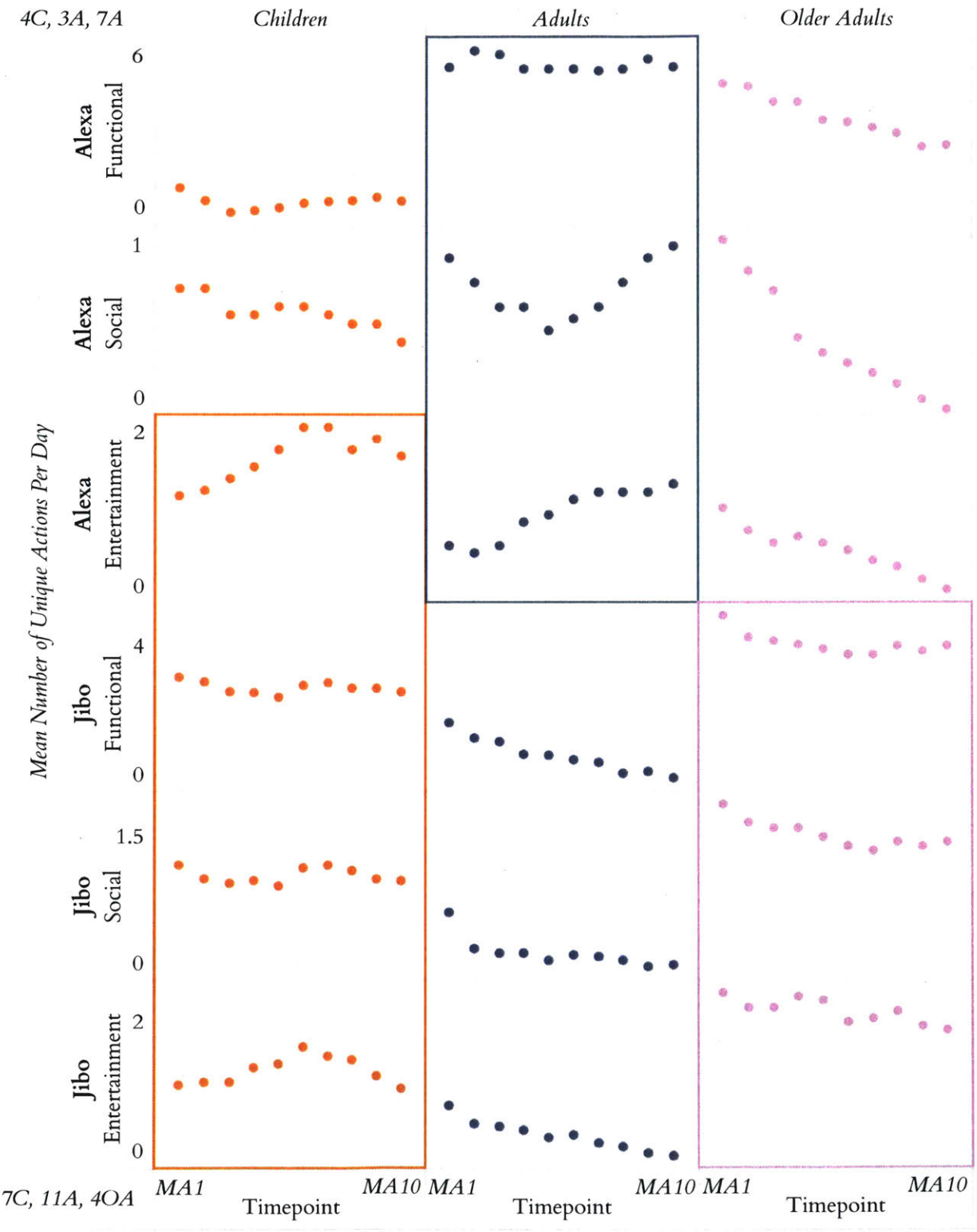


Figure 7.4: 5-day moving averages for mean number of unique actions by generation, action type, and agent type.

agents and action categories, but had slightly higher stability in their usage of Jibo. These observations for children and older adults, however, was impacted by C-11 and OA-11 who had consistently high usage each day. While longer-term data and larger sample sizes are necessary to find conclusive results, it is interesting to note differences in patterns of use across populations. For example, adults appeared to explore agents by prioritizing functional capabilities, later integrating social- and entertainment-related roles. This often came in the form of actions such as playing music, games, jokes, and asking questions about the agent's personality. On the other hand, older adults exhibited very similar decreasing trends across all agent action categories, suggesting a lack of disambiguation in use. Children, though particular about preferences, showed more consistency in their exploration over the 2 weeks.

THE IMPACT OF PRIOR EXPOSURE

To evaluate whether these differences in patterns of use were influenced by prior exposure to an agent, usage of participants who previously had a voice-based agent (Alexa n =11) was compared to those whose first introduction to living with the technology was with Jibo. The results of this, broken down by agent action category are shown in Table 7.5 below along with results from corresponding Welch T-Tests between the two conditions.

As shown, participants who previously owned an Amazon Alexa device and were assigned Jibo for the study (n=11, 4C, 6A, 1OA) had significantly lower usage across all action categories in

Table 7.5: Comparison of usage of participants who previously had a voice-based agent (Alexa n =11, no Google Home) vs. those whose first introduction to living with the technology was with Jibo

Agent Action Category	**Number of Unique Actions Per Day**		**Welch T-Test (Between Conditions)**
	Jibo Only (n=13)	*Alexa + Jibo (n=11)*	
Functional	$\mu = 3.49$, $\sigma = 5.20$	$\mu = 1.41$ $\sigma = 1.91$	$T(335) = -5.48$ $P^* = 1.09e^{-07}$
Social	$\mu = 1.13$ $\sigma = 1.82$	$\mu = 0.33$ $\sigma = 0.65$	$T(335) = -5.00$ $p^* = 1.08e^{-06}$
Entertainment	$\mu = 1.36$ $\sigma = 1.85$	$\mu = 0.51$ $\sigma = 0.86$	$T(335) = -5.49$ $p^* = 1.06e^{-07}$
Total	$\mu = 5.98$ $\sigma = 8.61$	$\mu = 2.26$ $\sigma = 2.84$	$T(335) = -5.55$ $p^* = 7.09e^{-08}$

comparison with participants whose first experience with the technology was with Jibo (n=13, 4C, 6A, 3OA). To further characterize these differences, 5-day moving averages were computed for both groups (Alexa + Jibo and Jibo respectively) and agent action category (functional, social, emotional), similar to those presented in Figure 7.4. Results of this are shown in Figure 7.5 below. As shown, participants who previous owned an Alexa device tended to have limited usage of Jibo across all categories and time points. In contrast, children and older adults whose first introduction to voice-based agents occurred with Jibo exhibited significantly higher and more stable usage of the technology. Adults showed minimal and decreasing usage in both groups.

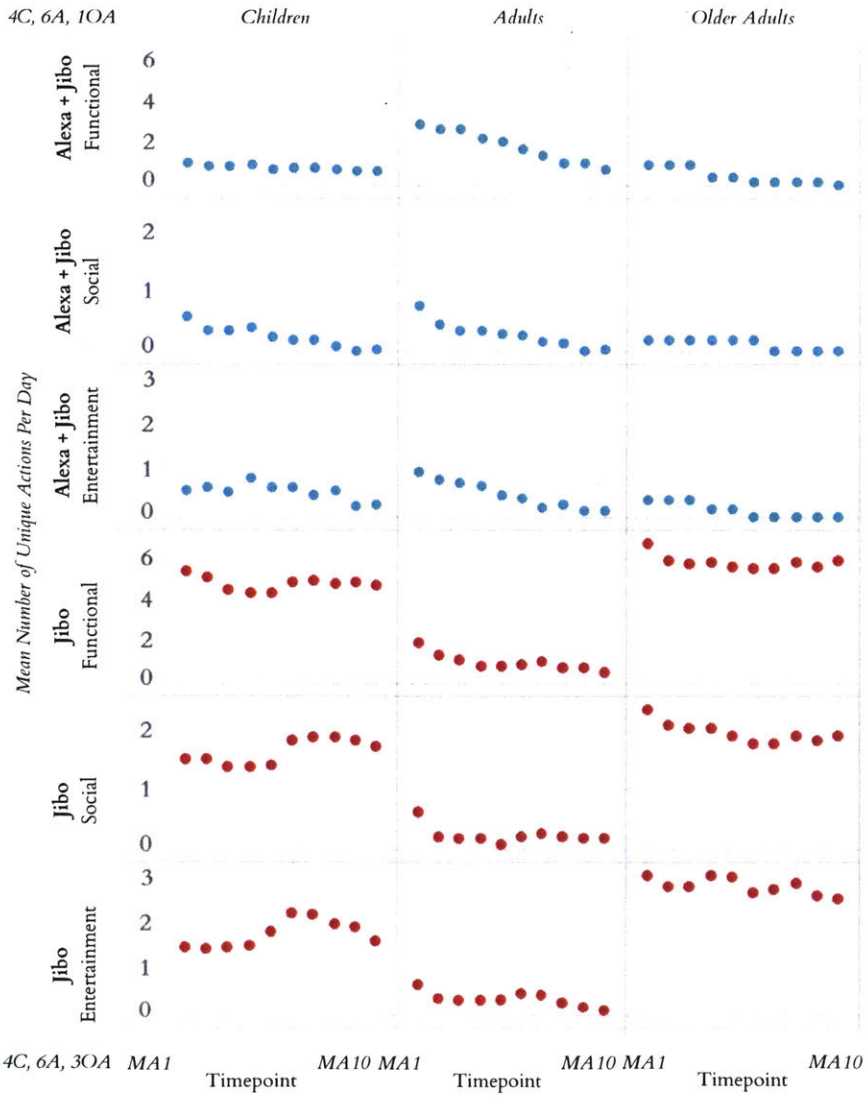


Figure 7.5: 5-day moving averages for mean number of unique actions by generation, action type, and agent type for participants who previously owned an Alexa and were given a Jibo for the study and those whose first experience with voice-based agents was with Jibo.

These findings may suggest that (1) participants with both agents saw them as distinct pieces of technology and used them each differently, (2) participants with prior exposure to agents were less curious about the technology or (3) participants with an Alexa preferred or were more used to the more functional, simple form. While further study is required to disambiguate reasons for this observation, it is helpful to note when designing for consolidation.

7.3 Personality

Another aspect of participants' experience with agents involved capturing perceptions of personality across the two agents. The goal of this analysis was to better identify (1) how participants perceived the two agents differently and (2) whether participants' perception of the agent's personality was driven by aspects of their own personality.

To capture this, the Ten Item Personality Measure (TIPI), a 10-item measure of the commonly-used Big Five personality dimensions was utilized. It is relevant to note that the TIPI measure utilizes 5 pairs of traits (each with a corresponding reverse trait), each corresponding to a Big 5 trait. Participant personalities were collected in addition to their perceptions of analogous traits defined for agents. Table 7.6 below provides a breakdown of the TIPI traits alongside defined agent-personality dimensions and Big 5 associations.

Table 7.6: TIPI personality traits, corresponding agent personality traits and relation to Big 5 traits.

#	TIPI Personality Trait	Agent Personality Trait	Trait Pair	Big 5 Trait
1	Extraverted, enthusiastic	Outgoing and engages me lots	2	*Extraversion*
2	Reserved, quiet	Quiet and keeps to itself	1	
3	Open to new experiences, complex	Always learning about me	4	*Openness to new experiences*
4	Conventional, uncreative	Simple in personality	3	
5	Sympathetic, warm	Sympathetic, warm	6	*Agreeableness*
6	Critical, quarrelsome	Opinionated and shares its thoughts	5	
7	Dependable, self-disciplined	Dependable and tries to help me	8	*Conscientiousness*
8	Disorganized, careless	Confused at times and may mess up	7	
9	Calm, emotionally stable	Consistent and predictable	10	*Emotional Stability/Reservedness*
10	Anxious, easily upset	Upset and worried if I'm not around	9	

Agent Personality Trait	Mean, SD, and Medians for Rating (Likert Scale 1-7) (7=strongly agree)		Mann-Whitney Tests (Between Agents)
	Alexa (n=19)	Jibo (n=25)	
Outgoing and engages me lots	$\mu = 3.11, \tilde{x} = 3.00$ $\sigma = 2.05$	$\mu = 4.64, \tilde{x} = 5.00$ $\sigma = 1.58$	$U = 125.00, p^* = 3.62e^{-03}$
Quiet and keeps to itself	$\mu = 5.84, \tilde{x} = 6.00$ $\sigma = 1.54$	$\mu = 3.36, \tilde{x} = 3.00$ $\sigma = 1.91$	$U = 76.50, p^* = 5.87e^{-05}$
Always learning about me	$\mu = 3.31, \tilde{x} = 3.00$ $\sigma = 2.06$	$\mu = 3.72, \tilde{x} = 3.00$ $\sigma = 1.97$	$U = 204.00, p = 0.21$
Simple in personality	$\mu = 5.11, \tilde{x} = 5.00$ $\sigma = 2.02$	$\mu = 4.32, \tilde{x} = 4.00$ $\sigma = 1.82$	$U = 168.50, p^* = 0.05$
Sympathetic, warm	$\mu = 3.52, \tilde{x} = 4.00$ $\sigma = 2.17$	$\mu = 4.12, \tilde{x} = 4.00$ $\sigma = 1.92$	$U = 199.50, p = 0.18$
Opinionated and shares its thoughts	$\mu = 2.89, \tilde{x} = 2.00$ $\sigma = 2.21$	$\mu = 3.16, \tilde{x} = 3.00$ $\sigma = 1.89$	$U = 209.50, p = 0.25$
Dependable and tries to help me	$\mu = 4.11, \tilde{x} = 4.00$ $\sigma = 2.02$	$\mu = 4.24, \tilde{x} = 5.00$ $\sigma = 1.83$	$U = 232.00, p = 0.45$
Confused at times and may mess up	$\mu = 4.05, \tilde{x} = 5.00$ $\sigma = 1.96$	$\mu = 5.64, \tilde{x} = 6.00$ $\sigma = 1.80$	$U = 117.50, p^* = 1.86e^{-03}$
Consistent and predictable	$\mu = 4.53, \tilde{x} = 6.00$ $\sigma = 2.44$	$\mu = 3.28, \tilde{x} = 3.00$ $\sigma = 1.77$	$U = 158.50, p^* = 0.03$
Upset and worried if I'm not around	$\mu = 1.11, \tilde{x} = 1.00$ $\sigma = 0.32$	$\mu = 2.72, \tilde{x} = 2.00$ $\sigma = 2.03$	$U = 127.00, p^* = 1.45e^{-03}$

Table 7.7: Comparison of participant perceptions of Alexa and Jibo across TIPI personality traits.

To identify differences in participants' perception of both agents' personalities, means, medians, and standard deviations were computed for each personality trait. A series of Mann-Whitney tests were conducted to evaluate significance of values for each personality trait (Table 7.7). Results indicated that participants perceived Jibo to be more outgoing ($\tilde{x}_J = 5.00$, $\tilde{x}_A = 3.00$, $U = 125.00$, $p^* = 3.62e^{-03}$), and conversely, less reserved ($\tilde{x}_J = 6.00$, $\tilde{x}_A = 3.00$, $U = 76.50$, $p^* = 5.87e^{-05}$) than Alexa. Participants perceived Alexa to be significantly more consistent and predictable ($\tilde{x}_J = 3.00$, $\tilde{x}_A = 6.00$, $U = 158.50$, $p^* = 0.03$) and consequently, slightly less confused ($\tilde{x}_J = 6.00$, $\tilde{x}_A = 5.00$, $U = 117.50$, $p^* = 1.86e^{-03}$). Alexa's personality was also perceived to be slightly more simple than Jibo's personality ($\tilde{x}_J = 4.00$, $\tilde{x}_A = 5.00$, $U = 168.50$, $p^* = 0.05$). Rather interestingly, there was higher agreement ($\mu_J = 2.72$, $\tilde{x}_J = 2.00$, $\mu_A = 1.11$, $\tilde{x}_A = 1.00$) and greater variability ($\sigma_J = 2.03$, $\sigma_A = 0.32$) in participants' perception of whether the agent would be "upset and worried if they were not around" for Jibo than Alexa ($U = 127.00$, $p^* = 1.45e^{-03}$).

These personality traits were then mapped to Big 5 personality traits as per the protocol described for the TIPI measure. The results of this are visualized in Figure 7.6 and detailed in Table 7.8 below. Participants perceived Jibo to be more extraverted ($\tilde{x}_J = 5.00$, $\tilde{x}_A = 2.50$) and open ($\tilde{x}_J = 3.50$, $\tilde{x}_A = 3.00$) while Alexa was perceived as more conscientious ($\tilde{x}_J = 4.50$, $\tilde{x}_A = 6.00$) and emotionally reserved ($\tilde{x}_J = 4.00$, $\tilde{x}_A = 3.25$).

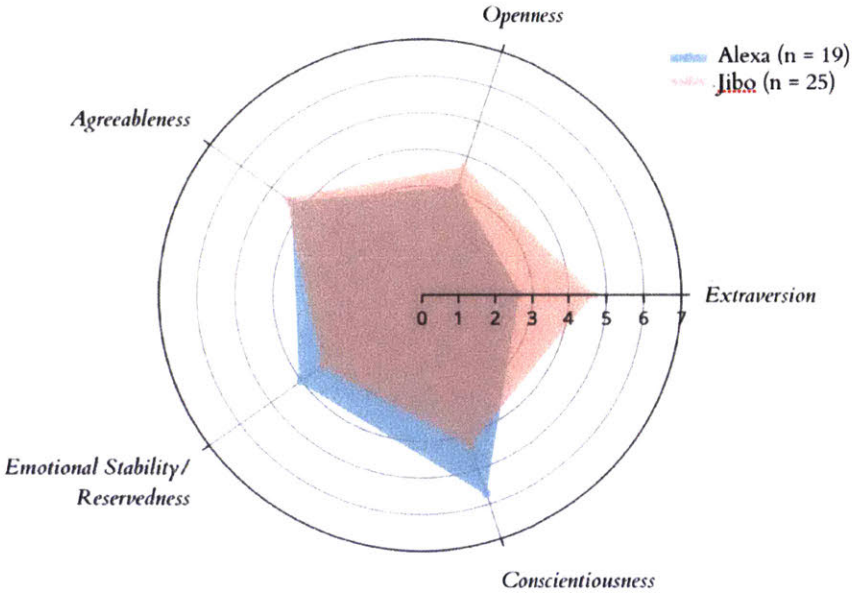


Figure 7.6: Radar chart of agent personalities along Big 5 Personality Traits.

Big 5 Trait	Jibo (n = 25)	Alexa (n = 19)	Mann-Whitney (Between Agents)
Extroversion	$\mu = 4.67$, $\tilde{x} = 5.00$, $\sigma = 1.11$	$\mu = 2.63$, $\tilde{x} = 2.50$, $\sigma = 1.52$	U = 149.50, $p^* = 2.30e\text{-}04$
Openness	$\mu = 3.68$, $\tilde{x} = 3.50$, $\sigma = 1.37$	$\mu = 3.11$, $\tilde{x} = 3.00$, $\sigma = 1.70$	U = 69.50, $p^* = 4.74e\text{-}07$
Conscientiousness	$\mu = 4.29$, $\tilde{x} = 4.50$, $\sigma = 1.40$	$\mu = 5.71$, $\tilde{x} = 6.00$, $\sigma = 1.25$	U = 222.50, $p^* = 0.01$
Emotional Stability/Reservedness	$\mu = 3.27$, $\tilde{x} = 3.25$, $\sigma = 1.25$	$\mu = 3.11$, $\tilde{x} = 4.00$, $\sigma = 1.31$	U = 198.50, $p^* = 3.96e\text{-}03$
Agreeableness	$\mu = 4.50$, $\tilde{x} = 4.00$, $\sigma = 1.17$	$\mu = 4.32$, $\tilde{x} = 4.00$, $\sigma = 0.95$	U = 259.00, $p = 0.06$

Table 7.8: Comparison of Alexa and Jibo along Big 5 Personality Traits. Mann-Whitney tests were conducted between agents to evaluate significance of differences.

Perceptions of agreeableness did not have significant differences between the two agents.

To dive deeper and identify whether participants' personalities influenced their perceptions of agent personalities, a series of 10 Pearson correlations were computed between participants' traits and participant' perceptions of their agent's traits for each agent. Results indicated no statistically significant correlations between participant personalities and their perceptions of Alexa's personality. However, perceptions of Jibo's personality traits did exhibit statistically significant results. Self and agent perceptions of extroversion were negatively correlated, indicating that more introverted participants tended to perceive Jibo as more extroverted (coefficient = -0.45, n = 25, $p^* = 0.02$) and vice versa. On the other hand, participants who perceived themselves to be more calm and emotionally stable tended to perceive Jibo as more consistent and predictable (coefficient = 0.44, n = 25, $p^* = 0.03$). Finally, there was weak evidence that participants who saw themselves as being open to new experiences and complex tended to perceive Jibo as learning less about them (coefficient = -0.34, n = 25, p = 0.09).

Differences in perceptions of agent personalities demonstrate several key ideas. Firstly, it was natural for people to project personality traits onto voice-based agents. Social expressivity in Jibo led to participants seeing the agent as more extraverted and open. However, it also resulted in lowered perceptions of conscientiousness and consistency, but greater projection of feelings towards. On the other hand, a less expressive agent such as Alexa was seen as more consistent, dependable, and stable, but also less personable. It is relevant to note that the perception of an agent's personality became a function of the perceiver's personality only when the agent was highly expressive. As such, it is important to consider that more variability in human perception may emerge when designing agents capable of exhibiting more expressive and social behaviors.

7.4 Artifacts of exploration

Throughout the agent home stay, participants shared notes, photos, videos, and "wish tokens" to capture their desires, reflections, and feedback. Some of these notes and artifacts are shown in the Figure 7.7, 7.8, and 7.9 below.

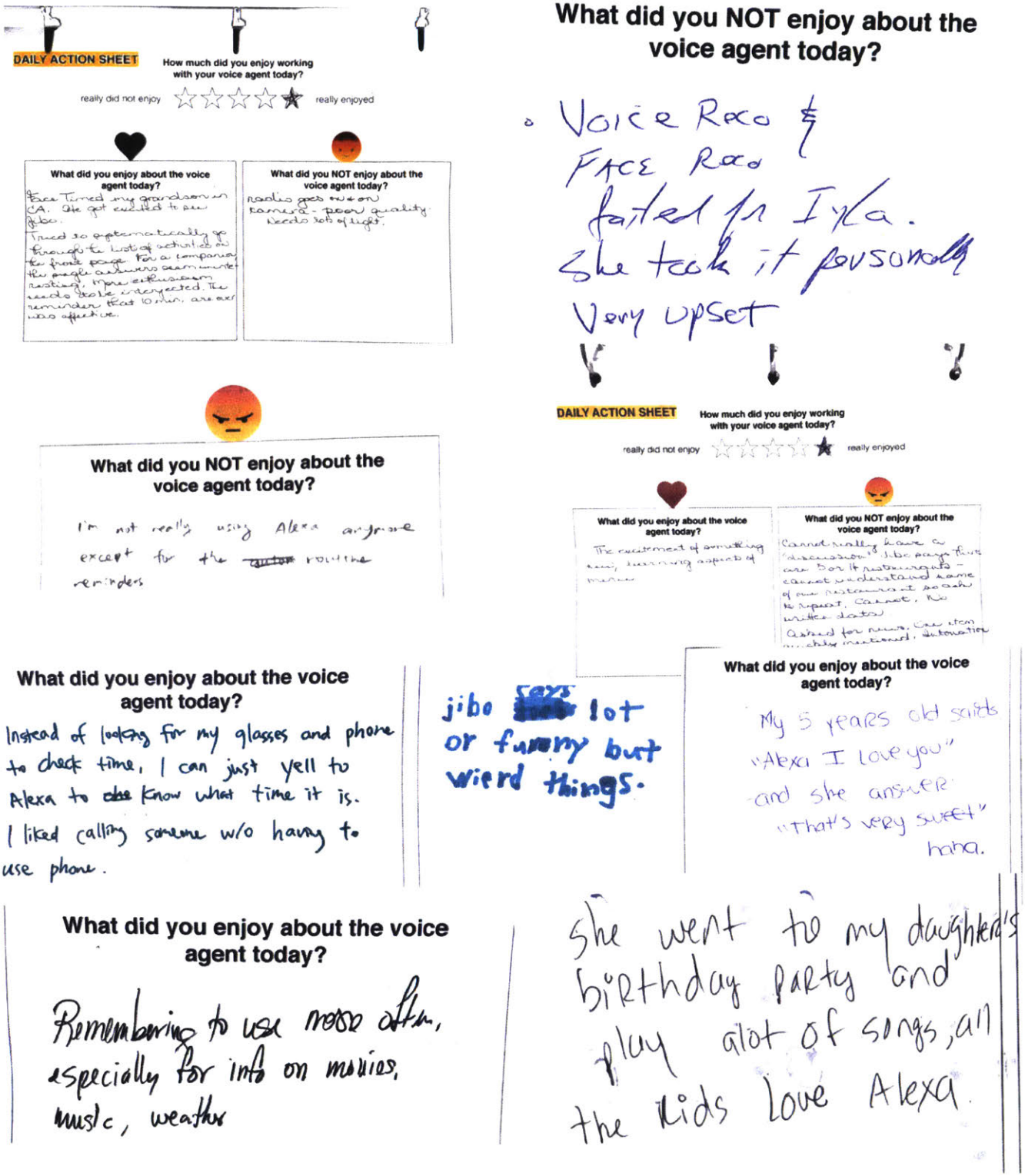


Figure 7.7: Examples of participant artifacts from home-stay.

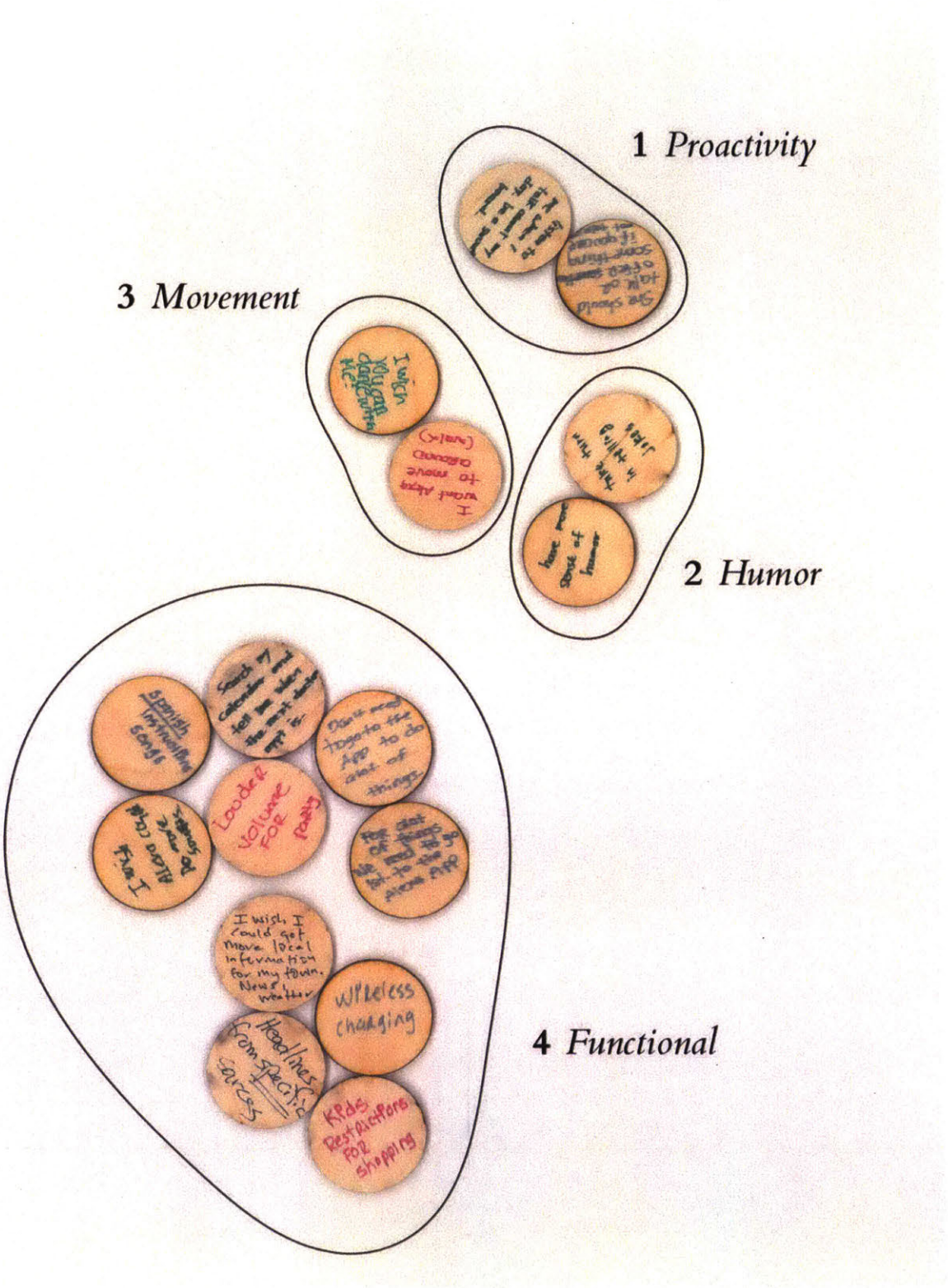


Figure 7.8: Wish tokens written by participants about Alexa, organized into general categories.

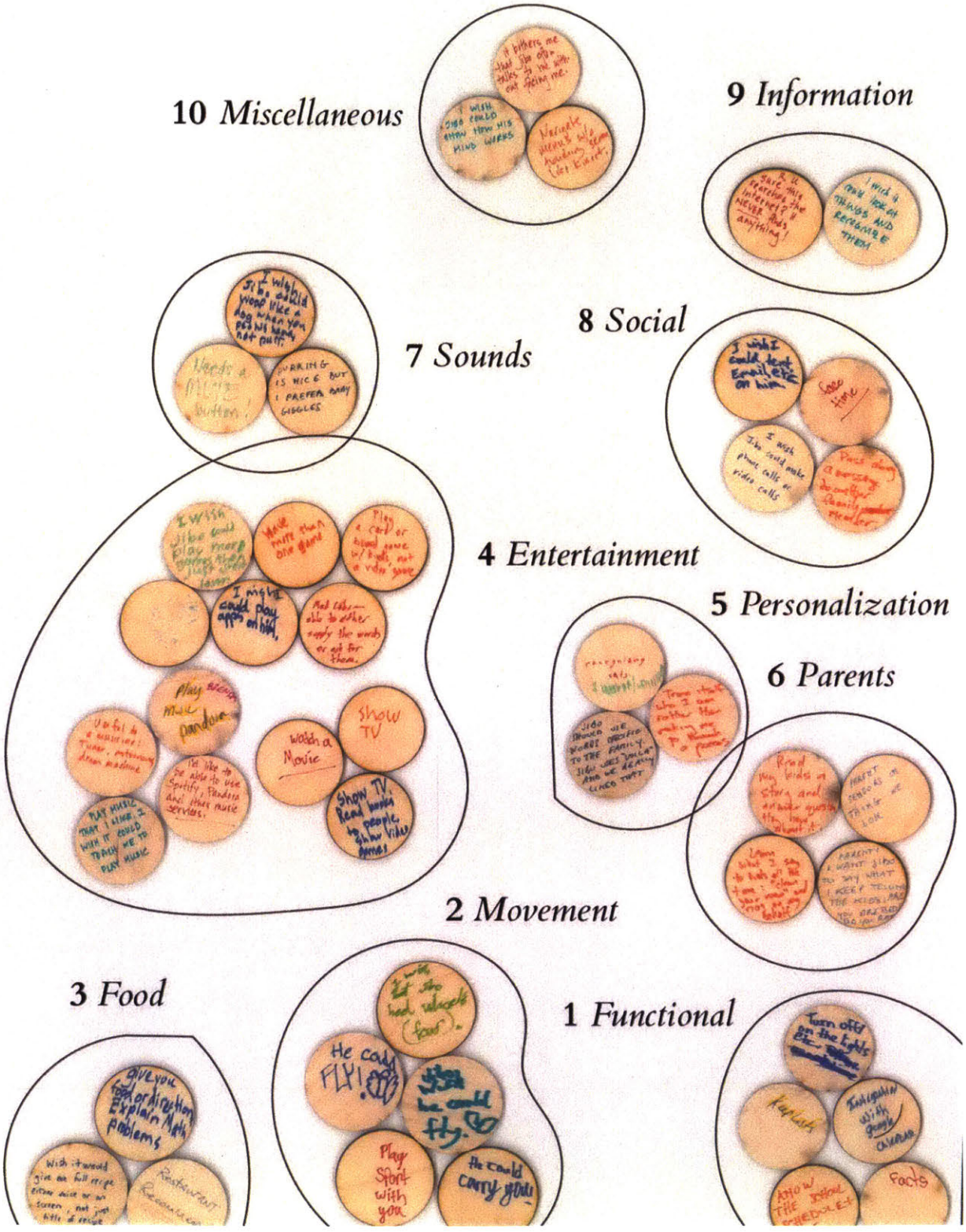


Figure 7.9: Wish tokens written by participants about Jibo, organized into general categories.

Notes shared by participants were well distributed across both agents and explored a range of thoughts including likes and dislikes, interesting moments, frustrations, and stories of the agents in their social environment (i.e. taking the agent to a birthday party or visits from friends). Unlike these notes, "wish tokens" for future agent capabilities were shared far more often by participants who had a Jibo (Figure 5.8) than Alexa (5.7). Wishes for Alexa tended to be highly functional and pragmatic in nature, with some desires for movement and humor. On the other hand, wishes for Jibo spanned a much wider spectrum—"cat face" recognition, flying and moving, ability to show multimedia content (i.e. movies, TV, games), and the ability to share how "his mind works" to name a few. Rather interestingly, the socially expressive nature of Jibo seemed to elicit more creativity, openness and expectations in participants' perspectives towards voice-based agents. While further exploration is required, these findings are highly relevant for designing behaviors of voice-based agents to fit the intended context.

7.5 *Insights*

The goal of this phase of study 2 was to understand the experiences of people living with voice-based agents in their home. This chapter explored (1) first day interactions, (2) longer-term patterns of use, (3) perceptions of personality, and (4) artifacts shared by participants. Below is a summary of the findings from the experiential component of the first study as relevant to the larger goal of this work.

(8) **USERS EXHIBIT PATTERNS IN TECHNOLOGY EXPLORATION**

Across generations, users broadly exhibited 3 patterns of exploration with the technology. One set of users were habitual in their usage of the agent, interacting with the agent for a few interactions everyday. The second group of users had a highly sporadic pattern in their use, remembering to interact with the agent at random intervals. Finally, a third group of users exhibited a decreasing trend over time, with usage stopping completely within the two weeks. It is relevant to consider these ideas to inform the design of technology behaviors for maintaining engagement. For example, by identifying a participant's pattern of use,

an agent could intelligently surface elements of surprise at relevant moments to re-capture the user's interest. Further exploration is required with longer-term deployments of agents in the home to fully assess the nature of these user patterns.

(9) GENERATIONS USE AGENTS DIFFERENTLY

Across the three generations, usage patterns and experiences of the technology varied significantly. Children exhibited the greatest stability in their use of both agents with entertainment-focused interactions. When the experience was more interactive (i.e. in the case of Jibo), children also used more functional aspects of the technology. Adults on the other hand placed far more importance on the functional aspects of the technology and desired a more "assistant-like" experience, showing much higher usage with Alexa than Jibo. Adults using Alexa sustained their use of functional interactions, slowly integrating social and entertainment interactions into their daily usage over time. Finally, older adults tended to evaluate the agent as a whole. While older adults' usage of functional actions was consistent across agents, social and emotional interactions were more common with Jibo. Further exploration of older adults' long-term use is required given the vast differences in aging that exist. These results are consistent with the analyses presented in study 1.

(10) THE TRADEOFFS OF PERSONALITY

Participants' perceptions of agent personalities differed significantly across the two agents. While it was natural for participants to describe the personality of the agents, more participants struggled with this task for Alexa's personality, sometimes stating that they were "*not sure she had one*". In contrast, participants projected more personal qualities onto Jibo. Jibo's dynamic persona and expressivity led participants to perceive the agent as more extraverted and open to new experiences, but also led to increased perceptions of negative qualities such as lack of conscientiousness and lower emotional reservedness. Conversely, Alexa was seen as more consistent, dependable, and stable, but less personable. Interestingly, when the agent was highly expressive like Jibo,

the perception of the agent's personality became dependent on the perceiver's personality, particularly around extraversion and emotional stability. These findings emphasize the importance and tradeoffs of persona development when designing interactions with voice-based agents.

(11) WITH EXPRESSIVITY COMES EXPECTATION

Participants tended to have broader and greater exploration of Jibo during the first day of interaction. Additionally, participants shared disproportionately more "wishes" for Jibo. These wishes were often more creative and expansive in nature. Together, these findings suggest that participants may have heightened expectations for voice-based agent technologies that integrate social-emotional behaviors.

PART III

8.0 Reflection: Analysis & Discussion

In the second phase of study 2, participants returned to the lab to reflect upon and share their experiences living with the agents. This reflection included a semi-structured interview and the card sorting activity to better understand how participants' perspectives had evolved. In this chapter, (1) the changes in agent action card choices are explored to identify trends across generations, action categories, and agent conditions, and (2) qualitative data coded from semi-structured interviews and card sorting is analyzed to understand participants' experiences living with the technology and identify surrounding themes and boundaries.

8.1 Quantitative Analysis

EVOLUTION OF CHOICES

To understand how living with agents affected overall placement of the 41 cards into the "yes", "neutral" and "no" buckets, the total proportion of cards per bucket was computed by study and generation. Study 1 is referred to as *pre* and study 2 is referred to as *post*. These results are summarized in Figure 8.1 for the 34 participants (8 children, 14 adults, 12 older adults) that provided card sorting results for both studies. Some participants (n = 7) were not able to complete both card sorting activities due to availability. As such, they have been excluded from the data presented to allow for effective comparison. Overall, all generations exhibited consistent placement of cards into the various buckets between pre and post. Overall, all generations increased the number of cards placed in the "yes" bucket

(pre_C = 58%, $post_C$ = 63%, pre_A = 58%, $post_A$ = 62%, pre_{OA} = 59%, $post_{OA}$ = 63%). Children and adults showed a decrease in the total proportion of cards placed in the "no bucket (pre_C = 21%, $post_C$ = 16%, pre_A = 21%, $post_A$ = 18%). Older adults, on the other hand, exhibited increases in the proportion of cards in the "yes" (pre_{OA} = 59%, $post_{OA}$ = 63%) and "no" buckets (pre_{OA} = 24%, $post_{OA}$ = 25%), but a decrease in the proportion of cards placed in "neutral" bucket (pre_{OA} = 16%, $post_{OA}$ = 12%).

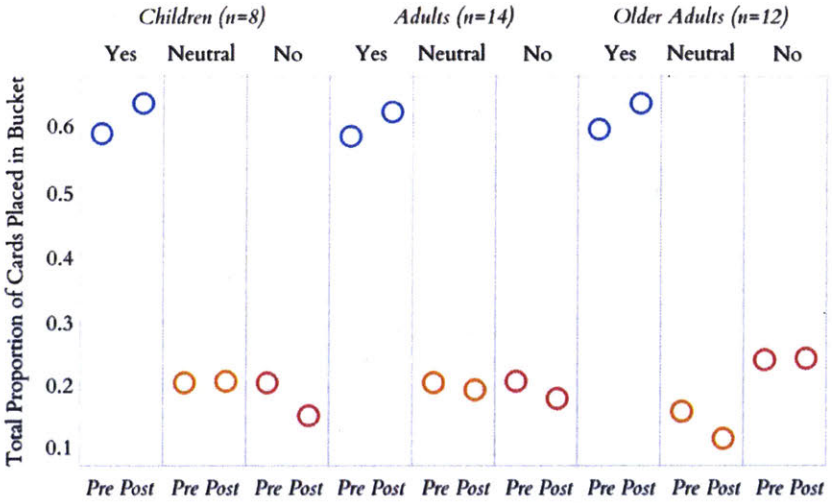


Figure 8.1: Changes in the total proportion of cards placed in "yes", "neutral" and "no" buckets by generation between the pre and post session.

To evaluate overall changes in agent action category choices, the proportion of cards placed in the "yes" bucket in the pre session and post session were computed by category. Corresponding Wilcoxon Signed-Rank tests were utilized to evaluate whether these differences in results between pre and post were statistically significant. Results are shown in Table 8.1.

Table 8.1: Mean, standard deviations, and medians for proportion of cards placed in the "yes" bucket by agent action category for the pre and post session. Corresponding Mann-Whitney tests evaluate significance of differences in pre and post.

	Pre **Proportion in "Yes"**			***Post*** **Proportion in "Yes"**			**Wilcoxon Signed-Rank Test Between Pre & Post**
Category	μ	$\tilde{x}$	σ	μ	$\tilde{x}$	σ	
Reminders	0.74	0.71	0.25	0.74	0.71	0.25	T = 252.50, p = 0.44
Information	0.62	0.64	0.27	0.66	0.71	0.29	T = 247.50, p = 0.39
Suggestions	0.42	0.40	0.26	0.37	0.25	0.32	T = 221.50, p = 0.19
Agent sharing something	0.59	0.60	0.33	0.66	0.80	0.29	T = 201.50, p = 0.10
Someone trying to reach	0.68	0.67	0.36	0.83	1.00	0.26	T = 99.00, $p^* = 6.81e^{-04}$
Someone sharing something	0.58	0.60	0.33	0.71	0.78	0.28	T = 201.00, p = 0.09

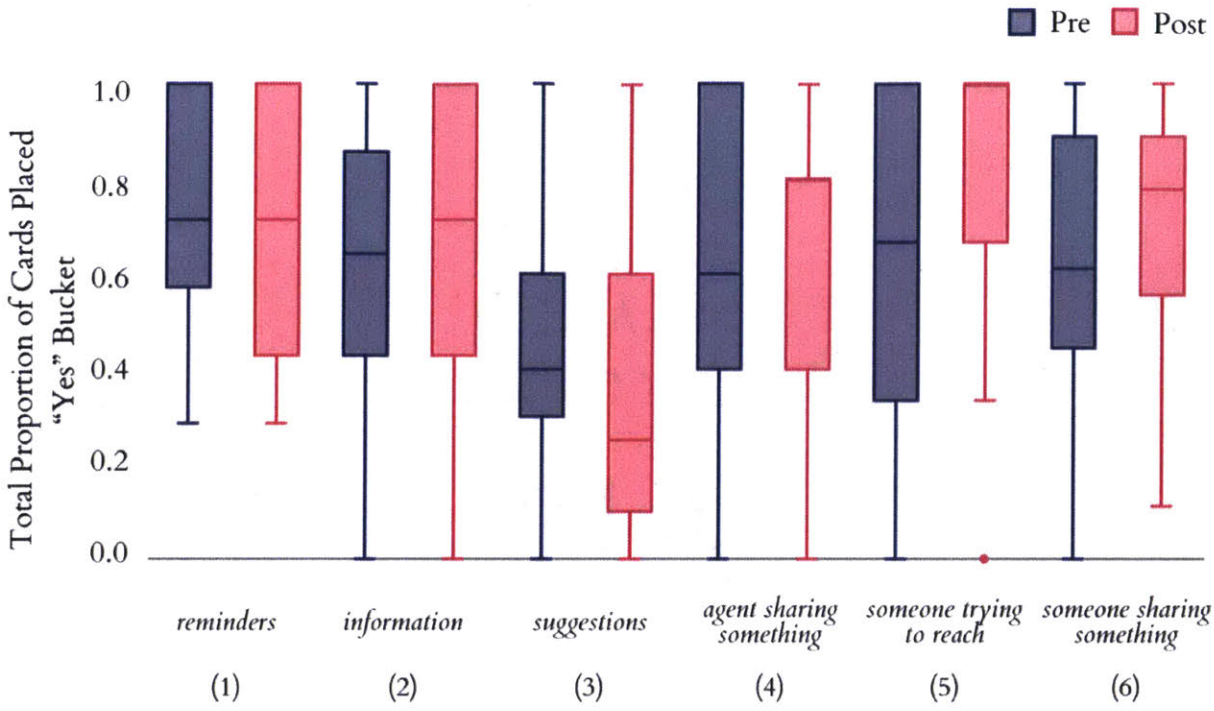


Figure 8.2: Boxplot of proportion of cards placed in the "yes" bucket across 34 participants in the pre and post session by agent action category

A corresponding boxplot characterizing distributions for pre and post by agent action category is shown in Figure 8.2. Distributions for reminders and information remained largely consistent, with slightly increases in spread. On the other hand, the proportion of cards placed in the "yes" bucket for suggestions decreased between pre and post ($\mu_{PRE} = 0.42$, $\tilde{x}_{PRE} = 0.40$, $\mu_{POST} = 0.37$, $\tilde{x}_{POST} = 0.25$).

Interestingly, all socially-driven categories—the agent sharing something ($\mu_{PRE} = 0.59$, $\tilde{x}_{PRE} = 0.60$, $\mu_{POST} = 0.66$, $\tilde{x}_{POST} = 0.80$) someone trying to reach the user through the agent ($\mu_{PRE} = 0.68$, $\tilde{x}_{PRE} = 0.67$, $\mu_{POST} = 0.83$, $\tilde{x}_{POST} = 1.00$), and someone sharing something through the agent ($\mu_{PRE} = 0.58$, $\tilde{x}_{PRE} = 0.60$, $\mu_{POST} = 0.71$, $\tilde{x}_{POST} = 0.78$)—exhibited significant increases in affinity and decreases in overall spread. Wilcoxon signed-rank tests between pre and post per category showed strong significance for someone trying to reach the user through the agent ($T = 99.00$, $p^* = 6.81e^{-04}$) and weak significances for the agent sharing something ($T = 201.50$, $p = 0.10$) and someone sharing something through the agent ($T = 201.50$, $p = 0.10$). While these results *do* indicate an overall evolution of participant perspectives after experiencing agents in the home, it is relevant to consider the nature of these differences across generations and agent types to identify drivers of change.

HOW GENERATIONS CHANGE DIFFERENTLY

To better understand these changes across agent action categories between pre and post, generational factors were considered. To characterize how agent action card choices

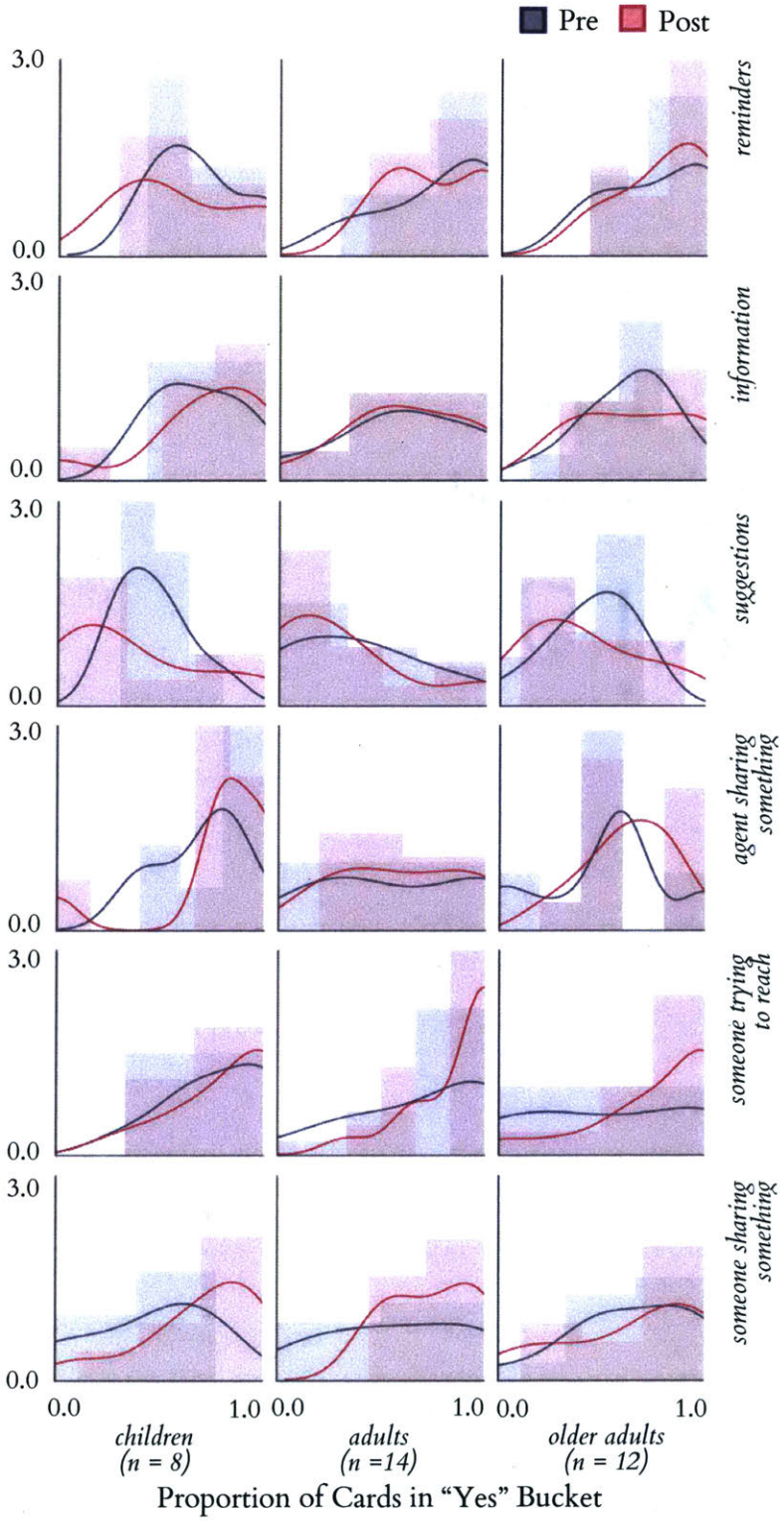


Figure 8.3: Histogram and distributions for the total proportion of cards placed in the "yes" bucket by category and generation for pre (blue) and post (pink).

evolved by generation, distributions for the proportion of cards placed in the "yes" bucket by generation and agent action category were plotted. Figure 8.3 shows these distributions for pre (blue) and post (pink). In addition, means, medians, and standard deviations were computed for pre, post, and the delta (post - pre) for each agent action category and generation. These results are shown in Table 8.2 below.

Table 8.2: Means, medians, and standard deviations for pre, post, and pre – post (delta) by agent action category and generation. Highlighted values indicate negative (pink) and positive (turquoise) changes above 3%.

Category	*Pre* Proportion in "Yes"			*Post* Proportion in "Yes"			*Post – Pre (Delta)* Proportion in "Yes"		
	μ	$\tilde{x}$	σ	μ	$\tilde{x}$	σ	μ	$\tilde{x}$	σ
Reminders									
Children (n = 8)	0.68	0.64	0.23	0.62	0.43	0.31	-0.05	-0.07	0.40
Adults (n = 14)	0.77	0.86	0.27	0.76	0.71	0.23	-0.01	0.00	0.20
Older Adults (n = 12)	0.76	0.79	0.25	0.79	0.86	0.22	+0.02	0.00	0.30
Information									
Children (n = 8)	0.68	0.64	0.21	0.70	0.79	0.33	+0.02	0.00	0.30
Adults (n = 14)	0.60	0.57	0.34	0.62	0.57	0.30	+0.02	0.00	0.30
Older Adults (n = 12)	0.62	0.71	0.23	0.67	0.71	0.29	+0.05	0.00	0.24
Suggestions									
Children (n = 8)	0.46	0.45	0.33	0.38	0.20	0.34	-0.09	-0.15	0.42
Adults (n = 14)	0.38	0.35	0.33	0.31	0.35	0.34	-0.06	-0.10	0.35
Older Adults (n = 12)	0.44	0.50	0.21	0.42	0.35	0.29	-0.02	0.00	0.21
Agent sharing something									
Children (n = 8)	0.70	0.80	0.21	0.77	0.80	0.33	+0.08	+0.10	0.30
Adults (n = 14)	0.60	0.80	0.38	0.61	0.80	0.32	+0.01	0.00	0.27
Older Adults (n = 12)	0.52	0.60	0.32	0.63	0.60	0.22	+0.12	+0.20	0.36
Someone trying to reach									
Children (n = 8)	0.79	0.83	0.25	0.83	1.00	0.25	+0.04	0.00	0.21
Adults (n = 14)	0.71	0.83	0.34	0.88	1.00	0.21	+0.17	0.00	0.36
Older Adults (n = 12)	0.56	0.50	0.43	0.78	1.00	0.33	+0.22	0.00	0.30
Someone sharing something									
Children (n = 8)	0.44	0.56	0.31	0.71	0.83	0.29	+0.26	+0.28	0.37
Adults (n = 14)	0.60	0.61	0.36	0.75	0.72	0.21	+0.17	+0.11	0.43
Older Adults (n = 12)	0.66	0.67	0.30	0.66	0.83	0.34	0.00	-0.06	0.48

To future capture these differences across generations, the proportion of participants that changed the number of cards (negative change, no change, positive change) in the "yes" bucket between pre and post was computed. These results are shown in the table below by generation and agent action category. Together, these analyses can be used to derive insight into how generations shifted perspectives after living with voice-based agents in their home.

Proportion of Participants that Changed # of Cards in "Yes" Bucket Between Pre & Post

Category		Type	neg	zero	pos
reminders	*n=8*	C	0.50	0.25	0.25
	n=14	A	0.29	0.43	0.29
	n=12	OA	0.25	0.33	0.42
information		C	0.25	0.50	0.25
		A	0.43	0.21	0.36
		OA	0.33	0.25	0.42
suggestions		C	0.63		0.38
		A	0.57	0.21	0.21
		OA	0.42	0.25	0.33
agent sharing something		C	0.25	0.25	0.50
		A	0.36	0.36	0.29
		OA	0.25	0.17	0.58
someone trying to reach		C	0.13	0.63	0.25
		A	0.14	0.43	0.43
		OA	0.08	0.50	0.42
someone sharing something		C	0.25		0.75
		A	0.29	0.21	0.50
		OA	0.50	0.08	0.42

Figure 8.4: Heat map showing the proportion of participants that changed the number of cards in the "yes" bucket between pre and post by generation and agent action category. Changes are characterized as negative, no change, and positive. Note that rows may not sum exactly to 1.0 due to rounding errors.

Across generations, more functional, or assistant-type functionalities such as reminders and information experienced the least change before and after experience living with agents. With respect to reminders, children exhibited an overall negative change ($\mu_{DELTA} = -0.05$, $\tilde{x}_{DELTA} = -0.07$), with half the children (negative change = 0.50, zero change = 0.25, positive change = 0.25) decreasing the proportion of reminder cards placed in the "yes" bucket. Adults and older adults also exhibited subtle changes in their preference for reminders, with adults decreasing slightly on average ($\mu_{DELTA} = -0.01$, $\tilde{x}_{DELTA} = 0.00$), and older adults increasing slightly on average ($\mu_{DELTA} = +0.02$, $\tilde{x}_{DELTA} = 0.00$). All generations showed an increase in their affinity towards the agent providing information, through the differences were subtle for adults ($\mu_{DELTA} = +0.02$, $\tilde{x}_{DELTA} = 0.00$), and children ($\mu_{DELTA} = +0.02$, $\tilde{x}_{DELTA} = 0.00$). Older adults exhibited a higher mean increase in their desire for information (positive change = 0.42, zero change = 0.25, negative change = 0.33) through the agent with a mean delta of +0.05.

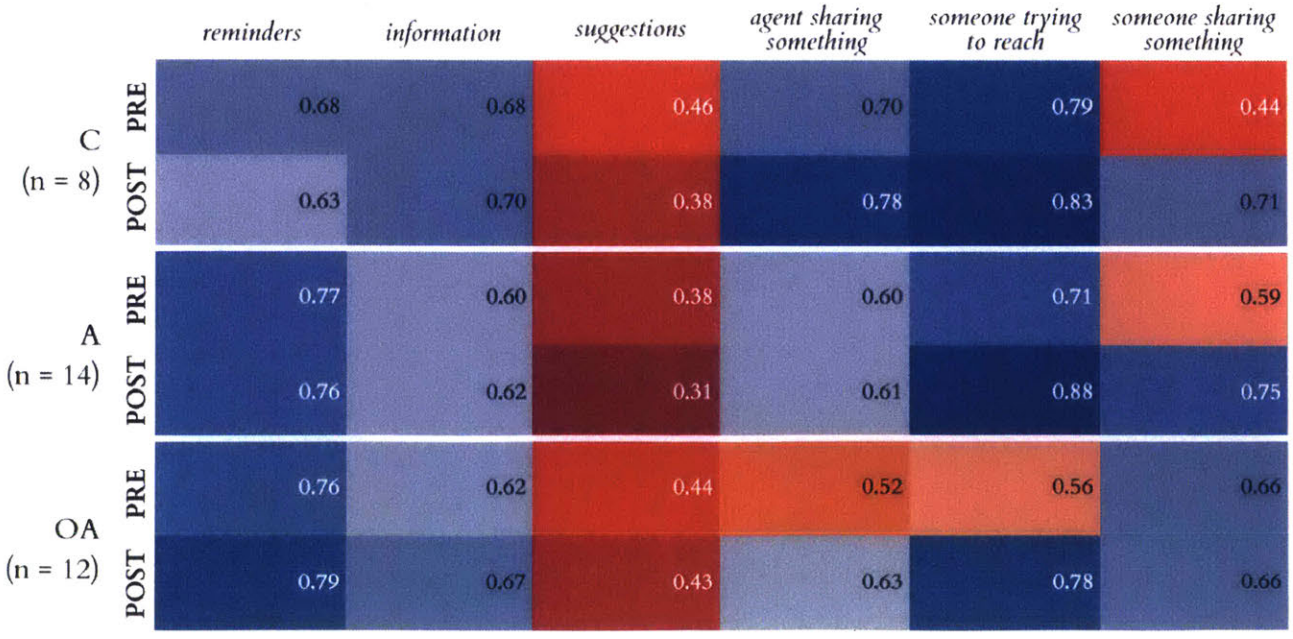
As shown in study 1, suggestions were often polarizing for all generations and exhibited the lowest affinity overall ($\mu_C = 0.38$, $\mu_A = 0.31$, $\mu_{OA} = 0.42$). This result remained consistent between pre and post, with mean desire of suggestions decreasing across all generations. These changes were most significant for children and adults. Rather interestingly, no children maintained equality in their desire for suggestions, with most showing an overall decrease (negative change = 0.63) and an overall mean delta of -0.09 and median delta of -0.15 between pre and post. Most adults became less open towards suggestions (negative change = 0.57) or stayed the same (zero change = 0.21), exhibiting a mean delta of -0.06 and a median delta of -0.10. Older adults exhibited a smaller overall decrease ($\mu_{PRE} = 0.44$, $\mu_{POST} = 0.42$).

The most noticeable changes between pre and post occurred in socially-driven categories across generations. All generations became more receptive to the agent itself sharing something with the user (e.g. music, artwork, thank you), with children and older adults exhibiting the most positive change. Children exhibited the highest affinity for this agent action category of all generations in the pre and post sessions, showing a mean increase of +0.08 and median increase of +0.10 in the proportion of cards placed in the "yes" bucket ($\mu_{PRE} = 0.70$, $\mu_{POST} = 0.77$). Older

adults also exhibited a substantial increase between pre and post from ($\mu_{PRE} = 0.52$, $\mu_{POST} = 0.63$, $\mu_{DELTA} = +0.12$, $\tilde{x}_{DELTA} = 0.20$). Adults did not show any substantial change between pre and post ($\mu_{DELTA} = +0.01$).

The last two agent action categories focused on the agent acting as a mediator for interaction with others, both through being a means for others to reach the user (e.g. phone call, status update) and as a means for others to share something (e.g. photo, video, song). With respect to the agent acting as a means to reach the user, all three generations exhibited increases in affinity towards the action (children: $\mu_{DELTA} = +0.04$, adults: $\mu_{DELTA} = +0.17$, older adults: $\mu_{DELTA} = +0.22$). On the other hand, while adults and children showed substantial overall increases in their preference for others sharing something through the agent (children: $\mu_{DELTA} = +0.26$, adults: $\mu_{DELTA} = +0.17$), older adults tended to bifurcate in their change of preference (negative change = 0.50, zero change = 0.08, positive change = 0.42), resulting in an overall minimal change ($\mu_{DELTA} = 0.00$, $\tilde{x}_{DELTA} = -0.06$). These results are summarized in the heat map in Figure 8.5 below for more context. As evident, the most significant changes occurred in the last three categories. It is relevant to recall that data shown in this heat map includes only participants who were present in both studies, and as such should only be used to interpret differences between pre and post.

Figure 8.5: Heat map comparing pre and post preferences between pre and post by generation and agent action category. Values indicate proportion of cards placed in the "yes" bucket.



Wilcoxon signed-rank tests were used to evaluate significance of these results between pre and post for each generation and agent action category. Results indicated strong significance for adults' and older adults' preference for the agent acting as a means to reach the user ($T_A = 19.00$, $p_A^* = 0.04$, $T_{OA} = 10.00$, $p_{OA}^* = 0.02$). Weak significance was also found for differences in children's affinity for someone sharing something through the agent between pre and post ($T_C = 5.50$, $p_C^* = 0.07$).

ACTIONS THAT CHANGE

To capture agent actions that experienced the most change in preference between pre and post, the number of participants who moved each agent action card from the "yes" to the "no" bucket and vice versa was computed. Bubble charts showing the top agent actions that were moved from "yes" to "no" and "no" to "yes" are shown in Figure 8.6 and Figure 8.7 respectively.

Suggestion-based agent actions such as a suggestion for reading or writing (e.g. new book etc.) (n = 7), suggestions to call (n = 7) or meet someone (n = 5), and suggestions to learn something new (n = 5) were most likely to be moved form "yes" to "no" buckets (n = 7). Drivers for these changes often stemmed from a desire for independence and autonomy. In addition, story-related actions were often moved from "yes" to "no", both for someone sharing a story (e.g. something about their day, a memory) through the agent (n = 7) and the agent itself sharing a story with the user (n = 4). Participants often described not wanting "*longer interactions*" and "*having specific tastes and preferences for content*" as justifications for these changes. Interestingly, despite the potential practical advantage, many participants moved to-do lists into the "no" bucket (n = 6). Reasons for this often stemmed from wanting to keep existing systems intact in the home (i.e. paper lists).

On the other hand, several socially-driven agent actions were moved from the "no" to the "yes" bucket. This was particularly true for *short-term* social interactions such as receiving a "hello" from someone (n = 7), the agent sharing a joke (n = 7), someone sharing a photo (n = 6), video (n = 6), status update (n = 6), and someone sending a "goodnight" (n = 5). Interestingly, several participants were more willing to receive a "thank you" from the agent (e.g. thank you for playing a game or interacting) than before (n

Figure 8.6: Yes to No
Bubble charts showing top cards and corresponding values for number of people that changed from "yes" to "no" buckets between pre and post.

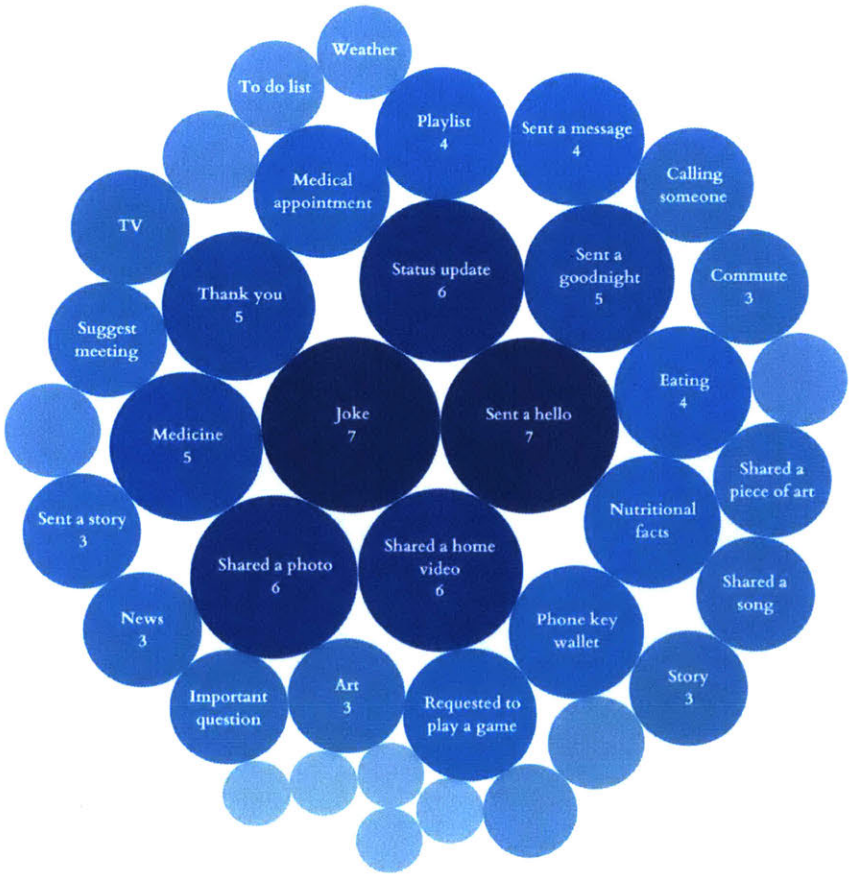


Figure 8.7: No to Yes
Bubble charts showing top cards and corresponding values for number of people that changed from "no" to "yes" buckets between pre and post.

the agent suggesting eating (i.e. time to eat, healthy food) (n = 4) and the agent sharing nutritional information (n = 4) also experienced an increase. This was often tied to participants beginning to envision the agent as a tool to support their wellbeing.

THE IMPACT OF LIVING WITH AGENTS

In addition to generational differences, it is important to consider the impact of experiencing highly distinct agents on participant preferences. To capture this, the pre and post proportions were compared for participants who lived with an Alexa and those who lived with Jibo by generation. Results of this analysis are shown in Table 8.3 below.

Children who experienced Alexa exhibited significantly higher increases in the proportion of cards placed in the "yes" bucket (μ_{DELTA} = +0.22). This difference was found to be statistically significant by a Wilcoxon signed-rank test between pre and post (T = 0.00, $p^{\star} = 2.22e^{-03}$). There was no significant change observed for children who experienced Jibo (μ_{DELTA} = -0.02). It is relevant to consider the smaller sample size for Alexa participants (n = 2) in comparison with Jibo (n = 6), as well as the higher variability in proportion for Jibo (σ_{JIBO} = 0.35, σ_{ALEXA} = 0.16) when interpreting these results. Adults, in contrast, did not exhibit any statistically significant differences between pre and post sessions for both agent groups. While adult participants with Alexa exhibited a slightly higher increase in the proportion of

Table 8.3: Pre, post and delta for proportion of cards placed in "yes" by agent and generation.

	Pre **Proportion in "Yes"**			***Post*** **Proportion in "Yes"**			***Post – Pre (Delta)*** **Proportion in "Yes"**			**Wilcoxon Signed-Rank Test Between Pre & Post**
Group	μ	$\tilde{x}$	σ	μ	$\tilde{x}$	σ	μ	$\tilde{x}$	σ	
Children										
Alexa (n = 2)	0.64	0.63	0.24	0.87	0.89	0.16	+0.22	+0.20	0.21	T = 0.00, $p^{\star} = 2.22e^{-03}$
Jibo (n = 6)	0.62	0.67	0.27	0.60	0.62	0.35	-0.02	0.00	0.36	T = 279.50, p = 0.40
Adults										
Alexa (n = 3)	0.43	0.41	0.26	0.54	0.44	0.28	+0.10	+0.05	0.30	T = 54.00, p = 0.17
Jibo (n = 11)	0.66	0.68	0.36	0.69	0.75	0.33	+0.03	0.00	0.34	T = 967.00, p = 0.38
Older Adults										
Alexa (n = 7)	0.54	0.56	0.30	0.59	0.60	0.28	+0.05	0.00	0.41	T = 389.50, p = 0.43
Jibo (n = 5)	0.66	0.71	0.32	0.75	0.87	0.31	+0.09	0.00	0.27	T = 99.00, $p^{\star} = 6.02e^{-03}$

cards placed in "yes", (Alexa: μ_{DELTA}= +0.10, Jibo: μ_{DELTA} = +0.03), these participants also had a significantly lower proportion of cards placed in "yes" during the pre-session (Alexa: μ_{PRE} = 0.43, Jibo: μ_{PRE} = 0.66). Given these differences, further exploration is required.

Older adults, in contrast, exhibited a higher increase in proportion of cards placed in "yes" after experiencing Jibo (Alexa: μ_{DELTA}= +0.05, Jibo: μ_{DELTA} = +0.09). A Wilcoxon signed-rank test found the differences in mean between pre and post to be statistically significant (T = 99.00, $p^{\star} = 6.02e^{-03}$).

To further evaluate the nature of these differences across generations and agent experiences, the proportion of cards placed in "yes" were computed by generation, agent, and agent action category. Results from this are shown Table 8.4 with major changes highlighted for ease. Purple highlights indicate opposing differences in the delta (i.e. one positive, one negative) between agents while turquoise highlights indicate same-sign differences. Corresponding results from Mann-Whitney tests conducted between agent conditions for each generation and agent action category are shown in Table 8.5. Given the exploratory nature of this work and likely influence of person-based factors such personality, demographics, technology affinity and more on results, it is important to interpret these results while considering such factors.

Reminders exhibited differences between agent conditions for children and older adults. This contrast was particularly strong for children. While children who lived with Alexa had a lower affinity for reminders during the pre-session (μ_{PRE} = 0.50) than children who lived with Jibo (μ_{PRE} = 0.74), participants with Alexa increased significantly (μ_{DELTA}= +0.50), in their desire for reminders while those with Jibo decreased (μ_{DELTA} = -0.24). This result was found to be statistically significant (U = 0.00, $p^{\star}$ = 0.03). However, given the smaller sample sizes and dissimilarity in pre values, further study is required. In contrast, older adults had similar affinity for reminders during the pre-session (Alexa: μ_{PRE} = 0.78, Jibo: μ_{PRE} = 0.74) but exhibited an opposing trend to children. Older adults who lived with Jibo tended to have a stronger desire for reminders in the post session on average (μ_{DELTA}= +0.11, σ_{DELTA} = 0.16) in comparison with those who lived with Alexa (μ_{DELTA} = -0.04, σ_{DELTA} = 0.37). These results were not found to be statistically significant, but do prompt

Category	*Pre* Proportion in "Yes"			*Post* Proportion in "Yes"			*Post – Pre (Delta)* Proportion in "Yes"		
	μ	$\tilde{x}$	σ	μ	$\tilde{x}$	σ	μ	$\tilde{x}$	σ
Reminders									
C, A (n = 2)	0.50	0.50	0.10	1.00	1.00	0.00	+0.50	+0.50	0.10
C, J (n = 6)	0.74	0.71	0.23	0.50	0.43	0.25	-0.24	-0.21	0.23
A, A (n = 3)	0.67	0.71	0.22	0.67	0.57	0.30	0.00	0.00	0.14
A, J (n = 11)	0.79	1.00	0.29	0.78	0.71	0.22	-0.01	0.00	0.22
OA, A (n = 7)	0.78	0.86	0.26	0.73	0.71	0.19	-0.04	0.00	0.37
OA, J (n = 5)	0.74	0.71	0.26	0.86	1.00	0.25	+0.11	0.00	0.16
Information									
C, A (n = 2)	0.93	0.93	0.10	0.93	0.93	0.10	0.00	0.00	0.00
C, J (n = 6)	0.60	0.57	0.17	0.62	0.64	0.35	+0.02	0.00	0.35
A, A (n = 3)	0.57	0.57	0.14	0.62	0.43	0.33	+0.05	-0.14	0.46
A, J (n = 11)	0.61	0.57	0.38	0.62	0.57	0.31	+0.01	0.00	0.27
OA, A (n = 7)	0.53	0.57	0.21	0.57	0.43	0.25	+0.04	0.00	0.29
OA, J (n = 5)	0.74	0.71	0.21	0.80	1.00	0.31	+0.06	0.00	0.16
Suggestions									
C, A (n = 2)	0.40	0.40	0.14	0.65	0.65	0.21	+0.25	+0.25	0.07
C, J (n = 6)	0.48	0.45	0.19	0.28	0.20	0.36	-0.20	-0.30	0.43
A, A (n = 3)	0.23	0.30	0.21	0.27	0.20	0.12	+0.03	+0.10	0.21
A, J (n = 11)	0.42	0.45	0.35	0.33	0.20	0.39	-0.09	-0.10	0.39
OA, A (n = 7)	0.39	0.50	0.23	0.34	0.30	0.14	-0.04	0.00	0.16
OA, J (n = 5)	0.52	0.50	0.18	0.54	0.70	0.41	+0.02	0.00	0.29
Agent sharing something									
C, A (n = 2)	0.70	0.70	0.14	0.90	0.90	0.14	+0.20	+0.20	0.00
C, J (n = 6)	0.70	0.80	0.24	0.73	0.80	0.37	+0.03	0.00	0.34
A, A (n = 3)	0.20	0.20	0.20	0.27	0.20	0.12	+0.07	0.00	0.31
A, J (n = 11)	0.71	1.00	0.35	0.71	0.80	0.29	0.00	0.00	0.27
OA, A (n = 7)	0.49	0.60	0.32	0.54	0.60	0.22	+0.06	0.00	0.43
OA, J (n = 5)	0.56	0.60	0.36	0.76	0.80	0.17	+0.20	+0.20	0.24
Someone trying to reach									
C, A (n = 2)	0.83	0.83	0.24	0.83	0.83	0.24	0.00	0.00	0.00
C, J (n = 6)	0.78	0.83	0.27	0.83	1.00	0.28	+0.06	0.00	0.25
A, A (n = 3)	0.56	0.33	0.38	0.78	0.67	0.19	+0.22	+0.33	0.51
A, J (n = 11)	0.76	1.00	0.34	0.91	1.00	0.33	+0.15	0.00	0.35
OA, A (n = 7)	0.48	0.33	0.42	0.81	1.00	0.26	+0.33	+0.67	0.58
OA, J (n = 5)	0.67	1.00	0.47	0.73	1.00	0.43	+0.07	0.00	0.15
Someone sharing something									
C, A (n = 2)	0.50	0.44	0.24	0.89	0.89	0.00	+0.39	+0.33	0.24
C, J (n = 6)	0.43	0.56	0.35	0.65	0.67	0.32	+0.22	+0.22	0.42
A, A (n = 3)	0.37	0.44	0.13	0.63	0.67	0.17	+0.26	+0.33	0.23
A, J (n = 11)	0.65	0.67	0.38	0.79	0.89	0.22	+0.14	0.00	0.48
OA, A (n = 7)	0.60	0.56	0.20	0.54	0.67	0.38	-0.06	-0.22	0.48
OA, J (n = 5)	0.73	0.89	0.42	0.82	0.89	0.23	+0.09	0.00	0.53

Table 8.4: Pre, post and delta for proportion of cards placed in "yes" by agent, generation and agent action category.

Category	Mann Whitney Tests Between Agents for Delta (Post – Pre)
Reminders	
C, A (n = 2) C, J (n = 6)	U = 0.00, p* = 0.03
A, A (n = 3) A, J (n = 11)	U = 16.50, p = 0.46
OA, A (n = 7) OA, J (n = 5)	U = 14.50, p = 0.34
Information	
C, A (n = 2) C, J (n = 6)	U = 6.00, p = 0.43
A, A (n = 3) A, J (n = 11)	U = 15.50, p = 0.47
OA, A (n = 7) OA, J (n = 5)	U = 17.00, p = 0.50
Suggestions	
C, A (n = 2) C, J (n = 6)	U = 2.00, p = 0.12
A, A (n = 3) A, J (n = 11)	U = 10.00, p = 0.17
OA, A (n = 7) OA, J (n = 5)	U = 16.00, p = 0.43
Agent sharing something	
C, A (n = 2) C, J (n = 6)	U = 2.50, p = 0.16
A, A (n = 3) A, J (n = 11)	U = 16.00, p = 0.50
OA, A (n = 7) OA, J (n = 5)	U = 12.00, p = 0.21
Someone trying to reach	
C, A (n = 2) C, J (n = 6)	U = 5.00, p = 0.42
A, A (n = 3) A, J (n = 11)	U = 15.50, p = 0.47
OA, A (n = 7) OA, J (n = 5)	U = 11.00, p = 0.15
Someone sharing something	
C, A (n = 2) C, J (n = 6)	U = 4.50, p = 0.37
A, A (n = 3) A, J (n = 11)	U = 10.50, p = 0.19
OA, A (n = 7) OA, J (n = 5)	U = 15.00, p = 0.37

Table 8.5: Mann Whitney tests between agent conditions (Alexa, Jibo) by generation and agent action category.

questions around the importance of form and interaction experience in desires for functionality.

Participant affinity for suggestions also exhibited differences between the agent conditions. Though children in both agent groups had similar desire for suggestions in the pre session (Alexa: $\mu_{PRE} = 0.40$, Jibo: $\mu_{PRE} = 0.48$), children who lived with Alexa exhibited a positive change on average in post ($\mu_{DELTA} = +0.25$), while children who lived with Jibo decreased significantly ($\mu_{DELTA} = -0.20$). It is relevant to note the high variability across children in the Jibo group ($\sigma_{DELTA} = 0.43$). This trend was echoed in adult participants, but with a lower change (Alexa: $\mu_{DELTA} = +0.03$, Jibo: $\mu_{DELTA} = -0.09$). Older adults exhibited an opposite, but subtle change (Alexa: $\mu_{DELTA} = -0.04$, Jibo: $\mu_{DELTA} = +0.02$). While there is not enough evidence to consider these results as statistically significant, these findings highlight the nuances and complexities of designing goal-directed behavior in the form of suggestions in agents.

This dichotomy between children's and older adults' experience living with the two agents was also observed when considering the agent sharing something with the user. Though children in both agent conditions had similar affinity for this agent action category in the pre session (Alexa: $\mu_{PRE} = 0.70$, Jibo: $\mu_{PRE} = 0.70$), children who experienced Alexa had a more significant increase in affinity in the post session (Alexa: $\mu_{DELTA} = +0.20$, Jibo: $\mu_{DELTA} = +0.03$). This trend was reversed in the case of older adults who became far more open to the agent sharing something after experiencing Jibo (Alexa: $\mu_{DELTA} = +0.06$, Jibo: $\mu_{DELTA} = +0.20$).

Agent actions related to the agent as a mediator of human-human connection (i.e. someone reaching the user through the agent and someone sharing something through the agent) tended to show increases across all agent conditions and generations. The one exception to this observation was older adults' affinity for someone sharing something from the user. Older adults who lived with Alexa tended to decrease in their desire for someone sharing something through the agent ($\mu_{DELTA} = -0.06$) in comparison with those who lived with Jibo ($\mu_{DELTA} = +0.09$). While these differences in agent conditions were not statistically significant, the variability observed across agent conditions and generations emphasizes the importance of interaction experiences and their ability to shape participant preferences.

8.2 Qualitative Analysis

FROM PRE TO POST

To capture the drivers for changes in participant preferences before and after living with agents, qualitative data from interview transcripts was analyzed. Presence values (i.e. proportion of participants who brought up the particular theme at least once) were determined for each theme defined in the coding scheme. A comparison of presence values between pre and post is provided in Table 8.6 by generation and positive or negative valence.

As evident from these breakdowns, functions the agent could help the user with (1.00), technology capabilities and use (1.00), and the agent's role (total = 1.00) continued to be the most common themes discussed by participants. After living with agents, themes surrounding impact (pre = 0.46, post = 0.57) and security and privacy (pre = 0.41, post = 0.91) became far more salient.

Participant discussions around functions the agent could help with remained largely consistent, with a few key differences. After experiencing the technology in their home, participants became far more focused on novelty (pre = 0.56, post = 1.00), both from a positive (1.00) and negative (0.70) perspective. In contrast, discussion surrounding the impact of the agent on participants' social environment became less salient (pre = 0.88, post = 0.76) and more negative (pre = 0.35, post = 0.58). Interestingly, though participants still discussed positive impacts of the technology on daily routines (pre = 0.95, post = 1.00), there was also a significant increase in negative discussion surrounding this theme (pre = 0.30, post = 0.94).

Another area of frequent discussion amongst participants was the capabilities and use of the technology. Participants were more likely to discuss the attractiveness (pre = 0.40, post = 0.94) or unattractiveness (pre = 0.72, post = 0.91) of the technology in the post session than pre. Despite this increase in perception of attractiveness, participants continued to perceive voice-based agents as more of a "nice-to-have" (positive = 0.85, negative = 0.48) rather than a "need-to-have" (positive = 0.39, negative = 0.88). Consolidation, personalization, and ease of technology also became more salient in the post session, with many participants articulating both positive and negative tradeoffs.

Category / Theme / Subtheme	Total	Positive	Negative	C	A	OA
	Proportion of Participants with Theme Present (n_{PRE} = 43, n_{POST} = 33) (Pre Value, Post Value)					
Functions Agent Can Help With	0.98, 1.00	0.98, 1.00	0.58, 0.93	0.92, 1.00	1.00, 1.00	1.00, 1.00
Daily Routines	0.95, 1.00	0.95, 1.00	0.30, 0.94	0.83, 1.00	1.00, 1.00	1.00, 1.00
Reminders	0.79, 0.91	0.77, 0.82	0.19, 0.27	0.58, 0.67	0.78, 1.00	0.91, 1.00
Information	0.79, 0.97	0.77, 0.94	0.07, 0.61	0.67, 0.83	0.44, 1.00	0.91, 1.00
Scheduling & Lists	0.40, 0.58	0.37, 0.55	0.07, 0.18	0.16, 0.17	0.55, 0.63	0.45, 0.75
Social Environment	0.88, 0.76	0.83, 0.69	0.35, 0.58	0.83, 0.67	0.78, 0.75	0.95, 0.75
Novelty	0.56, 1.00	0.56, 1.00	0.14, 0.70	0.66, 1.00	0.77, 1.00	0.41, 1.00
Health & Wellness	0.49, 0.64	0.44, 0.58	0.05, 0.15	0.17, 0.67	0.22, 0.56	0.77, 0.75
Technology Capabilities & Use	0.98, 1.00	0.79, 0.97	0.89, 1.00	0.83, 1.00	1.00, 1.00	1.00, 1.00
Attractiveness of Technology	0.79, 0.97	0.40, 0.94	0.72, 0.91	0.50, 0.89	1.00, 1.00	0.86, 1.00
Need-to-have	0.74, 0.91	0.28, 0.39	0.70, 0.88	0.42, 0.78	1.00, 0.94	0.82, 1.00
Nice-to-have	0.28, 0.91	0.21, 0.85	0.09, 0.48	0.08, 0.67	0.11, 1.00	0.45, 1.00
Consolidation of Technology	0.74, 0.91	0.58, 0.73	0.40, 0.64	0.50, 0.67	0.67, 1.00	0.91, 1.00
Personalization	0.58, 0.88	0.51, 0.73	0.23, 0.55	0.42, 0.67	0.89, 0.94	0.55, 1.00
Ease of Technology	0.26, 0.85	0.19, 0.61	0.09, 0.82	0.00, 0.67	0.00, 0.94	0.50, 0.88
Role	0.86, 1.00	0.67, 0.97	0.56, 0.94	0.75, 1.00	1.00, 1.00	0.86, 1.00
Autonomy	0.77, 0.82	-	-	0.58, 0.56	0.89, 0.88	0.82, 1.00
Person	0.69, 0.79	-	-	0.33, 0.56	0.89, 0.63	0.77, 1.00
Agent	0.65, 0.45	-	-	0.42, 0.33	0.78, 0.63	0.72, 0.63
Proactive Behavior	0.74, 0.88	0.74, 0.70	0.49, 0.61	0.50, 0.67	0.89, 1.00	0.82, 0.88
What	0.51, 0.55	0.47, 0.52	0.21, 0.21	0.17, 0.33	0.78, 0.69	0.59, 0.50
When	0.52, 0.58	0.40, 0.45	0.16, 0.27	0.33, 0.33	0.67, 0.81	0.50, 0.38
How	0.49, 0.27	0.37, 0.21	0.28, 0.12	0.08, 0.11	0.44, 0.31	0.68, 0.38
Who	0.21, 0.03	0.21, 0.03	0.00, 0.00	0.25, 0.00	0.33, 0.06	0.14, 0.00
Feelings Towards	0.05, 0.67	0.00, 0.48	0.05, 0.42	0.00, 0.33	0.00, 0.88	0.09, 0.63
Perceived Relationship	0.53, 0.70	-	-	0.33, 0.44	0.67, 0.75	0.59, 0.88
Assistant	0.32, 0.42	-	-	0.08, 0.00	0.44, 0.63	0.41, 0.50
Social Companion	0.28, 0.36	-	-	0.17, 0.11	0.33, 0.38	0.32, 0.63
Family or Friend	0.12, 0.15	-	-	0.17, 0.33	0.00, 0.06	0.14, 0.13
Child	0.02, 0.15	-	-	0.00, 0.00	0.00, 0.31	0.05, 0.00
Pet	0.02, 0.15	-	-	0.00, 0.11	0.00, 0.13	0.05, 0.35
Attributing Human Qualities	0.42, 0.97	0.28, 0.88	0.16, 0.82	0.33, 0.89	0.44, 1.00	0.45, 1.00
Presence in the Home	0.21, 0.91	0.07, 0.79	0.17, 0.76	0.00, 0.89	0.22, 1.00	0.32, 0.75
Location	0.02, 0.48	0.02, 0.42	0.02, 0.18	0.00, 0.18	0.00, 0.63	0.05, 0.38

Impact	0.46, 0.57	0.35, 0.37	0.21, 0.22	0.25, 0.22	0.44, 0.69	0.59, 0.75
Motivation	0.37, 0.33	0.33, 0.27	0.07, 0.15	0.25, 0.00	0.44, 0.44	0.41, 0.50
Accountability	0.26, 0.27	-	-	0.08, 0.11	0.22, 0.31	0.36, 0.38
For Self	0.21, 0.18	-	-	0.08, 0.11	0.11, 0.19	0.32, 0.38
For Others	0.09, 0.18	-	-	0.00, 0.00	0.22, 0.31	0.09, 0.00
Laziness	0.21, 0.21	0.07, 0.15	0.21, 0.09	0.00, 0.11	0.11, 0.12	0.36, 0.50
Security & Privacy	0.41, 0.91	0.14, 0.47	0.30, 0.74	0.50, 0.78	0.55, 1.00	0.32, 0.88
Interaction Experience	0.37, 0.82	0.11, 0.33	0.26, 0.67	0.50, 0.78	0.44, 0.81	0.27, 0.88
Data Collection	0.09, 0.15	0.02, 0.12	0.07, 0.12	0.00, 0.00	0.22, 0.31	0.09, 0.00

Table 8.6: Summary of theme presence in transcripts between pre and post. Results are provided overall, by sentiment, and by generation.

Exploration of the agent's role was another core theme in all participants' discussions (1.00). In comparison with the pre-session, participants were less likely to consider the agent's autonomy (pre = 0.65, post = 0.45) and more likely to consider their own autonomy in the agent-person relationship (pre = 0.69, post = 0.79). Interestingly, discussion around personal autonomy increased most with older adults (pre = 0.77, post = 1.00) and children (pre = 0.33, post = 0.56), but decreased in adults (pre = 0.89, post = 0.63). Mixed feelings towards proactive behavior (positive = 0.48, negative = 0.42) were also expressed by many participants (pre = 0.05, post = 0.67). Children, adults and older adults were highly distinct in their perspective on the perceived relationship with the agent. In the post session, fewer children described the agent as a social companion (pre = 0.17, post = 0.11), but far more children saw the agent as a family member or friend (pre = 0.17, post = 0.33). In contrast, more adults described the agent as an assistant (pre = 0.44, post = 0.63) or social companion (pre = 0.33, post = 0.38). Older adults were also most likely to describe the agent as a social companion (pre = 0.32, post = 0.63) or an assistant (pre = 0.41, post = 0.50). The experience of living with agents also led to greater attribution of human qualities to the agent (pre = 0.42, post = 0.97) and significant discussion around location and presence in the home.

While security and privacy was not a salient feature previously, its importance in participant discussion increased significantly in the post session (pre = 0.41, post = 0.91). It is interesting to note that discussions around security and privacy disproportionately focused on the interaction experience (0.82) rather than data collection (0.15). This trend was observed

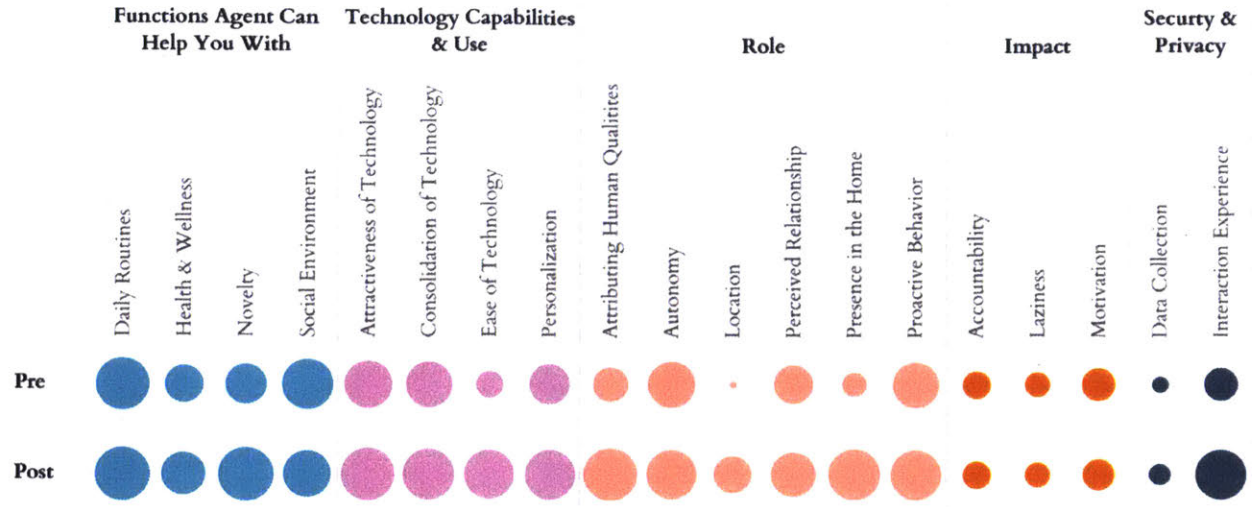


Figure 8.8: Bubble chart depicting relative presence of themes in pre and post. The size of the circle indicates proportion of participants who discussed each theme or subtheme.

across generations, with older adults and children not even discussing data collection once. Overall sentiments of participants discussing the interaction experience and its relation to security and privacy were largely negative (0.67), though some positive sentiments were captured (0.33). More participants also spoke about the impact of the technology (pre = 0.46, post = 0.57), particularly amongst adults (pre = 0.44, post = 0.69) and older adults (pre = 0.59, post = 0.75).

A visual summary of the differences in presence of themes between pre and post is provided in Figure 8.8 below. The size of the circle indicates the proportion of people who discussed each particular theme or subtheme. Data points are colored by the corresponding theme category.

HOW AGENTS AFFECT PERSPECTIVES

To better understand how the interaction dynamics of the agent impacted participant perspectives, presence of themes was determined by agent condition and generation. These values are summarized in Table 8.7 below. It is important to recall that themes were coded from transcripts of the entire post-session, including both the debrief and card sorting activity. As such, differences in theme presence across agent conditions should not be interpreted as *only* stemming from direct dialogue concerning the agent itself.

Participants who lived with an Alexa tended to discuss more about the core functionalities such as reminders (Alexa: 1.00,

Category / Theme / Subtheme	Proportion of Participants with Theme Present ($n = 33$, $n_{ALEXA} = 12$, $n_{JIBO} = 21$) (Jibo Value, Alexa Value)					
	Total (J, A)	Positive (J, A)	Negative (J, A)	C J = 6, A = 3	A J = 12, A = 4	OA J = 3, A = 5
Functions Agent Can Help With	1.00, 1.00	1.00. 1.00	0.95, 0.91	1.00, 1.00	1.00, 1.00	1.00, 1.00
Daily Routines	1.00, 1.00	1.00, 1.00	0.95, 0.91	1.00, 1.00	1.00, 1.00	1.00, 1.00
Reminders	0.86, 1.00	0.71, 1.00	0.38, 0,08	0.50, 1.00	1.00, 1.00	1.00, 1.00
Information	0.95, 1.00	0.90, 1.00	0.67, 0.50	0.83, 0.83	1.00, 1.00	1.00, 1.00
Scheduling & Lists	0.43, 0.83	0.38, 0.83	0.24, 0.08	0.17, 0.17	0.50, 1.00	0.67, 0.80
Social Environment	0.71, 0.83	0.66, 0.75	0.52, 0.67	0.67, 0.67	0.75, 0.75	0.67, 1.00
Novelty	1.00, 1.00	1.00, 1.00	0.76, 0.58	1.00. 1.00	1.00, 1.00	1.00, 1.00
Health & Wellness	0.62, 0.67	0.62, 0.50	0.14, 0.17	0.67, 0.67	0.67, 0.25	0.33, 1.00
Technology Capabilities & Use	1.00, 1.00	0.95, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00
Attractiveness of Technology	0.95, 1.00	0.90, 1.00	0.86, 1.00	0.83, 1.00	1.00, 1.00	1.00, 1.00
Need-to-have	0.86, 1.00	0.33, 0.50	0.81, 1.00	0.67, 1.00	1.00, 1.00	1.00, 1.00
Nice-to-have	0.95, 0.83	0.90, 0.75	0.57, 0.33	0.83, 0.33	1.00, 1.00	1.00, 1.00
Consolidation of Technology	0.86, 1.00	0.67, 0.83	0.57, 0.75	0.50, 1.00	1.00, 1.00	1.00, 1.00
Personalization	0.90, 0.83	0.76, 0.67	0.57, 0.50	0.83, 0.33	0.92, 1.00	1.00, 1.00
Ease of Technology	0.85, 0.83	0.57, 0.67	0.81, 0.83	0.67, 0.67	1.00, 0.75	0.67, 1.00
Role	1.00, 1.00	0.95, 1.00	0.95, 0.91	1.00, 1.00	1.00, 1.00	1.00, 1.00
Autonomy	0.81, 0.83	-	-	0.50, 0.67	0.92, 0.75	1.00, 1.00
Person	0.76, 0.83	-	-	0.50, 0.67	0.83, 0.75	1.00, 1.00
Agent	0.62, 0.17	-	-	0.50, 0.00	0.58, 0.00	1.00, 0.40
Proactive Behavior	0.90, 0.83	0.81, 0.50	0.71, 0.42	0.83, 0.33	1.00, 1.00	0.67, 1.00
What	0.62, 0.42	0.57, 0.42	0.23, 0.17	0.50, 0.00	0.67, 0.75	0.67, 0.40
When	0.71, 0.33	0.62, 0.17	0.33, 0.17	0.33, 0.33	0.91, 0.50	0.67, 0.20
How	0.33, 0.17	0.24, 0.17	0.19, 0.00	0.17, 0.00	0.33, 0.25	0.67, 0.20
Who	0.00, 0.08	0.00, 0.08	0.00, 0.00	0.00, 0.00	0.00. 0.25	0.00, 0.00
Feelings Towards	0.71, 0.58	0.57, 0.33	0.43, 0.42	0.50, 0.00	0.92, 0.75	0.33, 0.80
Perceived Relationship	0.76, 0.58	-	-	0.67, 0.00	0.75, 0.75	1.00, 0.80
Assistant	0.38, 0.50	-	-	0.00, 0.00	0.58, 0.75	0.33, 0.60
Social Companion	0.43, 0.25	-	-	0.17, 0.00	0.42, 0.25	1.00, 0.40
Family or Friend	0.14, 0.17	-	-	0.50, 0.00	0.00. 0.25	0.00, 0.20
Child	0.19, 0.08	-	-	0.00, 0.00	0.33, 0.25	0.00, 0.00
Pet	0.19, 0.08	-	-	0.17, 0.00	0.16, 0.00	0.33, 0.20
Attributing Human Qualities	0.95, 1.00	0.90, 0.83	0.90, 0.67	0.83, 1.00	1.00, 1.00	1.00, 1.00
Presence in the Home	0.90, 0.92	0.76, 0.83	0.86, 0.58	0.83, 1.00	1.00, 1.00	0.67, 0.80
Location	0.57, 0.33	0.48, 0.33	0.19, 0.16	0.50, 0.00	0.58, 0.75	0.67, 0.20

Impact	0.62, 0.50	0.38, 0.36	0.19, 0.27	0.17, 0.33	0.75, 0.50	1.00, 0.60
Motivation	0.33, 0.33	0.29, 0.25	0.14, 0.17	0.00, 0.00	0.50, 0.25	0.33, 0.60
Accountability	0.33, 0.17	-	-	0.17, 0.00	0.33, 0.25	0.67, 0.20
For Self	0.19, 0.17	-	-	0.17, 0.00	0.17, 0.25	0.67, 0.20
For Others	0.24, 0.08	-	-	0.00, 0.00	0.33, 0.25	0.00, 0.00
Laziness	0.24, 0.17	0.14, 0.17	0.10, 0.08	0.00, 0.33	0.17, 0.00	1.00, 0.20
Security & Privacy	0.95, 0.83	0.43, 0.55	0.90, 0.45	0.83, 0.67	1.00, 1.00	1.00, 0.80
Interaction Experience	0.81, 0.83	0.24, 0.50	0.76, 0.50	0.83, 0.67	0.75, 1.00	1.00, 0.80
Data Collection	0.23, 0.00	0.19, 0.00	0.19, 0.00	0.00, 0.00	0.42, 0.00	0.00, 0.00

Jibo: 0.86), information (Alexa: 1.00, Jibo: 0.95), and scheduling and lists (Alexa: 0.83, Jibo: 0.43). Positive sentiments surrounding these core functionalities were also more common in the case of participants who lived with Alexa, largely stemming from the greater capabilities currently offered by the device. Interestingly, discussion around health and wellness tended to be more common and positive with Jibo, suggesting that socially expressive behaviors may be important for agents that support wellbeing.

Technology capabilities and use were often discussed by participants, both for Jibo (1.00) and Alexa (1.00). While the dominant sentiment around the technology as a *need-to-have* was negative (Jibo: 0.81, Alexa: 1.00), participants were more positive about the technology as a *nice-to-have* (Jibo: 0.90, Alexa: 0.75). Consolidation of technology, personalization, and ease of technology were complex and nuanced spaces across agent conditions and participants, with a multitude of positive and negative sentiments. It is interesting to note that children discussed personalization with regards to Jibo far more often than with Alexa (Jibo: 0.83, Alexa: 0.33). This trend was reversed in adults (Jibo: 0.92, Alexa: 1.00).

The theme category with the greatest disagreement across agent conditions was the role of the agent. Discussions around agent autonomy in the person-agent relationship were disproportionately more common when participants who lived with Jibo (Jibo: 0.62, Alexa: 0.17). This trend was observed across all generations with children (Jibo: 0.50, Alexa: 0.00), adults (Jibo: 0.58, Alexa: 0.00), and older adults (Jibo: 1.00, Alexa: 0.40). Building on this, participants with Jibo also tended to talk a bit

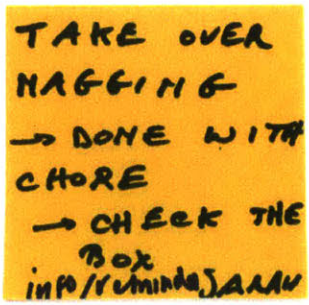
more about proactivity (Jibo: 0.90, Alexa: 0.83), particularly with regards to *what* (Jibo: 0.62, Alexa: 0.42) and *when* (Jibo: 0.71, Alexa: 0.33). Relational aspects of the technology were also more salient for participants who lived with Jibo (Jibo: 0.76, Alexa: 0.58). Jibo was more often discussed as a social companion (0.43), pet (0.19) or child (0.19), while Alexa was predominately seen as an assistant (0.50) and occasionally a social companion (0.25). Interestingly, no children that lived with Alexa discussed any thoughts about the perceived relationship (0.00) with the agent while those who lived with Jibo found this to be highly relevant (0.67).

Discussion surrounding impact were relatively uncommon compared to other theme categories across both agent conditions (Jibo: 0.62, Alexa: 0.50). All generations that lived with Jibo tended to discuss more ideas around accountability (Jibo: 0.33, Alexa: 0.17). Older adults who lived with Alexa discussed more ideas around motivation (Jibo: 0.33, Alexa: 0.60) and were less concerned about laziness (Jibo: 1.00, Alexa: 0.20).

Concerns around security and privacy were present for both agents (Jibo: 0.95, Alexa: 0.83). These concerns largely arose from the interaction experience itself. While participants who lived with Alexa were balanced in their expression of positive and negative sentiments (positive: 0.50, negative: 0.50), those who lived with Jibo tended to voice more concerns around security and privacy (positive: 0.24, negative: 0.76).

FUNCTIONS THE AGENT CAN HELP WITH

DAILY LIFE



Many participants described the integration of the agent into their daily routines. For families, particularly those with young children, the agent often became a useful feature in their morning habits: "*You will see in the little check box pages there are certain functions they were right on almost daily. Like I would ask Alexa about how the weather in the morning*". This use of the technology in this context often stemmed from busy schedules, the ease of hands-free voice access, and needing to "*get the kids out the door*". In this vein, some parents—particularly those who lived with Jibo—even described wanting the agent to "nag" their children on their way out the door in the morning: *"I want Jibo to keep saying what I tell the kids. Like are you dressed? Ten times. Did you eat? Like a zillion times."* Other parents also described

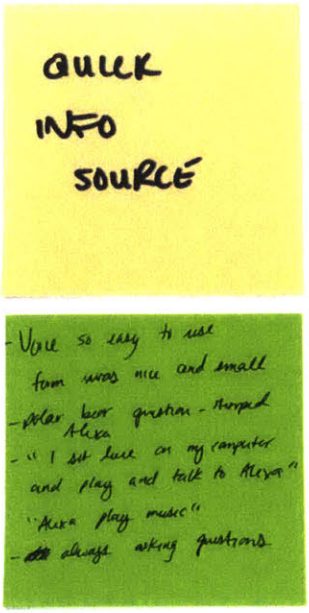
futuristic applications of the agents for context: *"The school buses have GPS systems on them so you can actually look online where the bus is. And I would like Jibo to be checked into the system and go the bus is on the corner of the street. You better start running".*

In addition to morning routines, many families with young children integrated the agent into their life for homework-related tasks. These applications often took the form of asking questions *("Like if we're studying something and we come to a questions. We'll be like "Alexa, what's the Great Wall?"*) and timers *("...part of her homework is reading for 20 minutes each day. So she would say, "Jibo, set a timer for 20 minutes". The one time she was so absorbed in the book, she didn't even notice it go off. She was like, it didn't work. So then she set a timer for another 20 minutes. So she did a lot of reading that day").* Several parents also described using the timers as a means for accountability to limit playtime and breaks.

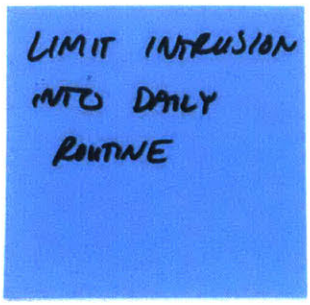
Beyond families, older adults also integrated the agent into their daily routines in creative ways for information access *("I used to help me with my crossword puzzles. So that was a good part of it.")* and reminders *("My experience was it was great for the more concrete rigid things. I loved that it reminded me to take my blood pressure every day. I never forgot my blood pressure a single day because it did just that...").* One older adult even described how he used Alexa as a 5AM alarm, asked her the weather and about scheduled medical appointments, and would often *"sit back on [his] computer and play and talk to Alexa"* on a daily basis.

Another area of common use and desire was to-do and grocery lists. Participants described liking the hands-free nature of the interaction, consolidation of information, and the ability to access the list when at the store using the phone application *("...that was one of the other better things about it because I could just walk around the kitchen and talk. You know and I didn't have to get a pencil paper, and then forget where I put it and go run after it. So that was another real plus".* Participants in larger families even saw the capacity for the agent to become a centralized source for all family members *("I actually have an app on my phone we keep for grocery list and somebody...it's...constantly...if you're the last one to use something and you don't write it down and then I think it would be good to just tell Jibo and not to have to just write it..."*—older adult).

While some participants were open to the idea of the technology assisting with to-do and grocery lists, others felt

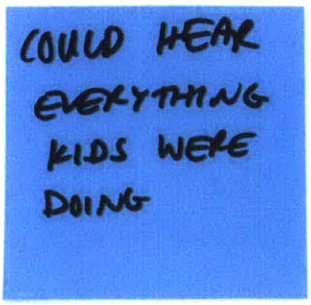


that voice-access to such content would be stressful and cumbersome (*"Hearing my to-do list? It would stress me out"*—adult). Other negative perspectives on the integration of agents into daily routines often stemmed from wanting to limit distractions, inconveniences surrounding the use of voice as an interface, and the technology not meeting expectations or needs *("It ended up feeling more like an incompetent assistant and then I just felt like every time it tried to initiate interactions with me I was annoyed at it"*—young adult, Parent: *"What about remembering what he says? What did you think about that?"* Child: *"Oh, yeah. He doesn't remember anything").* Participants often articulated higher expectations for Jibo, asking about functionalities as a "given"—*"But, I don't even understand why it can't do certain things with it. It's almost like they told it that it can't do it. It's not allowed to. Like an example, recipes. Why can't it do recipes?"*



SOCIAL ENVIRONMENT

For parents, the agent often became a "toy" for children. Several parents articulated a desire for the agent to entertain their children during busy moments—*"Like Jibo had the screen now and they have pictures, so entertain her for ten minutes while I'm doing something, that'd be great".* In fact, one child described talking to Alexa for over an hour, asking multiple questions and *"talking about video games and stuff…".* The appeal of this application of the technology often stemmed from the fact that the agent was *"not a TV"* and "*something that the family could sort of enjoy together*". Some families described interacting with the technology together as part of their social environment (*"I love the personality and we ask a lot of things like: OK, Alexa are you married? And she answered something super funny but I can't remember. And we'd love that. We laugh a lot and ask also just to have fun and she's super funny. So yeah, it's like was a good for funny"*—parent and kids). These ideas were often brought up alongside concerns about parental controls for the device and boundaries of role *("Yeah, no, thank you to stories. Mommy and daddy read stories to you so you don't need Alexa to do that for you. It's some family time").*

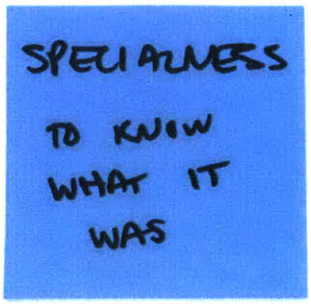


Several participants, across generations, also saw the potential for the agent to aide in interactions with external friends and family. Children described playing with the agent during playdates while adults and older adults even utilized it as a

tool during social gatherings (*"...we were trying to play a game and we no longer had the directions, so like "oh, Alexa, how to you play this game?"*—adult). Several older adults also emphasized the use of the technology as a means to help manage social relationships (*"yeah, definitely prompting you to get people together...so the companion is more of a social secretary"*) and connect with others (*"Oh, especially the people who never even seen such a thing or heard of such a thing. 'Cause I hang out with a lot of people who are not too up and coming in the modern world. But when my kids came over most of them knew all about it. All my grandchildren were far away, and when they hear about it from their parents or from me they are stunned."*). In fact, a few participants even brought the agent along to birthday parties and family gatherings.

NOVELTY

Novelty was another core part of the initial experience with agents. Many participants discussed initial fascination with the technology (*"Who would know what that little disk does if I hadn't found out by all my chats with you first? There it was, and every so often I'd say, "Play me something from Gershwin"; And it would be very nice, have some lovely music, and I thought of other people Harry Belafonte; it was lovely. It was wonderful. It was like having a pal there in the room."*—older adult). Many participants described regular use of, and enjoyment of Jibo's proactive "word of the day" interaction (*"I kind of liked that he would kind of start having conversations like word of the day"*—adult). Such interactions provided participants with a break in the daily routine and an opportunity to learn something new (*"Of course, the songs...the trivia...I love the trivia. It's like you learn something every day"*—adult, "*There were a lot of strange things that were actually hilarious"*—child). Many participants also described adapting the technology by interacting with features in unique ways such as trying to "stump the agent", creating stop-motion animations using Jibo's camera and more. This behavior tended to be more common with families who lived with Jibo than those who lived with Alexa.



After a few weeks however, participants also described a decrease in this novelty experience (*"I think three weeks was enough to go through the whole period where the novelty is really exciting. Then after that we overused a handful of the functions on*

it."—adult). Participants who were most unimpressed by the novelty aspect of the technology were young adults *("I think millennials have very high standards and expectations for technology in terms of functionality, so I feel like maybe older generations will appreciate the novelty of a social agent, whereas I think for us, we're used to things working and always wanting the best functionality. So I think that really shapes the experience").*

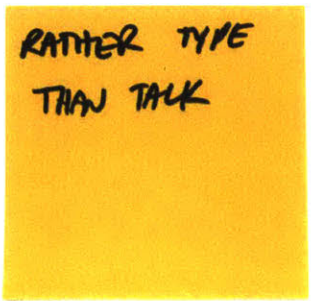
TECHNOLOGY CAPABILITIES & USE

ATTRACTIVENESS

The sentiment around whether participants felt they *needed* the technology remained consistent with feelings before living with the agents (negative: 1.00, positive: 0.50). Many participants, particularly adults, highlighted that most of the functionalities offered by the agents were easier to access in other technologies and that they didn't feel the agent enabled any change in lifestyle *("Why do I need this when I have my computer, or I can ask my phone?"*—adult, *"...it wasn't particularly useful, and I didn't feel strong about it. I don't think any of this would change my life in any way"*—adult). Positive sentiments from participants surrounding the attractiveness of the technology as a *nice-to-have* often stemmed from its ability to break the routine and provide a way to manage day-to-day tasks *("...learn something new. It reminded me of the word-of- the-day thing. In small doses, I think it can be a cool hobby"*—adult).

CONSOLIDATION

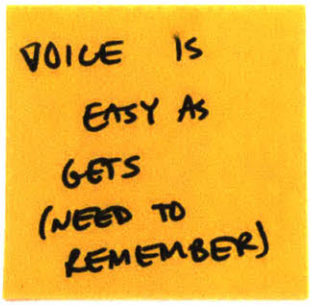
Building on perspectives towards the attractiveness of the technology, most participants saw the agents as a standalone device *("Not just that but I don't want to start thinking of my agent as just another dumb phone. Try to keep the devices as a seperate purpose almost."*—adult). All generations expressed reservations towards the agent integrating capabilities of the phone *("...I finally learned to use the cell phone. I use it on the road. I use it when I'm out and about. So I'm pretty comfortable with it. I don't think Alexa could take over that function. I don't see that"*—older adult, *"Because my phone can do all of this. That's why so I don't really know why..."*—child). Participants also highlighted their frustration with other forms of consolidation and their non-desire for agents to exist as part of that ecosystem *("I don't know if*



you have an Apple product…but sometimes like a phone call will come in on my IPad and my phone at the same time and its pretty annoying, so its assuming its done well"—adult).

While some participants were more critical about the attractiveness of the technology, others tended to be open to agents as "another thing" in their ecosystem *("…these are things I could do with my cell phone, so it would be neat if I could do it with this other system").* This perspective was common amongst older adults and parents with young kids who saw the hands-free and stationary nature of the technology as a means to access phone capabilities while busy with other tasks—*"I really don't like cellphones. And so, if it would be another landline, I would be happy with that. I know where it is and I'm not using it. I think the people around me would be happy about that too because no one can get in touch with me. And if like there would be a phone call on Jibo and I would just be walking around, I think I would like that".*

EASE OF USE



The most common frustration experienced by participants with respect to ease of use was the use of the wake word. The inconvenience of having to repeat the wake word seemed to conflict with the hands-free, convenient promise of the technology. Participants voiced frustrations on this matter across both agents, but often struggled more with the addition of the "hey" when interacting with Jibo *("But I did find that the "Hey Jibo" thing was another deterrent to that, because if you say like, "Hey Jibo, set a 30-minute timer" and it doesn't hear you, you then have to say like, "Hey Jibo" again, whereas if Alexa doesn't understand you, you just repeat "Alexa, set a timer"; So I think like the "Hey" thing"*—adult, *"I kind of wished you could just I don't know somehow engage it without always have to say the name"*—older adult). Several participants discussed enjoying the addition of the "hey" initially as a more personable greeting, but felt that using it each time felt unnatural *("Cause you wouldn't say like, Hey [name]. Hey [name]. Hey [name]"*—young adult).

Another area of difficulty articulated by participants was the challenge of discovery (*…sort of a mixed blessing with the problem of discovery. Like if you don't know what all of his functions are, but some of them can only be activated by very specific phrases. There's probably stuff in there we never would've found, even after years. But one of them, it volunteered for us. It just asked "do you want to play*

the word of the day?" Okay, so that became another one of our regular habits"—adult). These challenges present an interesting opportunity for technologists to consider designing *for* discovery and proactive surfacing of capability. In addition to frustrations with initiating interactions, participants also described the challenges of voice as an interface in group interactions *("So wait, so we are all talking at the same time. So just wait. That's something that I ended up putting in the jar. Or that it may be like primary boss or something. For example, if everyone talks at the same time, you are going to obey to mommy. And then to dad and then to the kids"*—adult).

On the other hand, positive sentiments towards ease of the technology often arose from the hands-free nature and resulting affordances *("Talking for me is much easier of course, because I have two kids or when I cooking I just talk or whatever"*—adult, *"The ease of technology just the interaction of not having to walk to a computer or pick up a computer"*—young adult).

PERSONALIZATION

Discussion around personalization was fairly expansive across generations and fell broadly into three core areas: interaction, language, and core functions.

Families often desired personalization of the agent's interaction and form to accommodate family taste *("We should put a tail on the back so when he dances..."*—child, *"There also was...we were discussing this morning whether it was an update or whether we found it later but it purrs. And, that was pretty popular but I think I would prefer baby giggles as a sound"*—adult). Several participants also discussed the interaction of the agent with their pets, often desiring "cat or dog face recognition" *("I don't remember Jibo ever saying out of the blue, "how was your cat today?" No, it never actually caught on to that about you, no matter how much you talked about cats and showed it cats"*—adult). Personality personalization was also important to older adults who wanted the agents to have more old-school, relatable personalities *("Oh well yeah. If I could have George Carlin or somebody like that, sure").* While participants with Alexa often commented on wanting more personality in the agent, discussions on more expansive aspects of personality personalization were more commonly observed with participants who lived with Jibo.

Another area of personalization discussed by participants was

surrounding the language of the agent. Some parents discussed the grown up nature of the agent dialogue (*"...but I thought it was very grown-up. Like she'd ask questions, but just you know, the language is very grown-up I'm like are you understanding that? So, wasn't necessarily geared towards a kid. Like it didn't differentiate itself..."*) while others wanted the agent to integrate elements of their family language style (*"What I also thought was funny was Jibo used the word Voila, which I use all the time. So I thought that was funny. I wondered if he could pick up certain works used in the family to like use that too. The first time Jibo said it, we were all like where did that come from?! So we thought that was funny"*).

Finally, a major aspect of personalization desired by participants related to the core functionalities of the device itself. After experiencing the current state of the technology, some participants became more skeptical about the potential for the technology to capture their very "*particular*" preferences. Many participants described wanting greater specificity in customization of the technology for their taste (*"I feel I'd enjoy, like the playlist. But I'd like to customize it a little bit more...like I usually ask about, play classical music. I want us to...like, because, specify, play classical music with piano. Like solo, or something like that"*—adult, *"I'm particular about my jokes, I don't think I'd want Alexa jokes. And again with art I'm not sure about that. I'm particular about that too, I'm not sure that Alexa would know, learn what I like in art"*—older adult). With a more socially expressive agent like Jibo, participants also expressed higher expectations for it to learn more about them (*... it didn't learn as much as it could have from us. Maybe my expectations were too high. But it didn't necessarily say "Hi Steve how is your day?" but it would pipe up and say maybe "hello". And I don't think I minded that. I wasn't shocked with it. I actually would sort of engage it then since it was sort of a way of talking. So from the social aspect I think that was fine"*—adult).

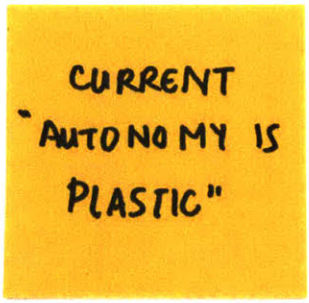
ROLE OF THE AGENT

AUTONOMY

Notions of agent and person autonomy were often explored by participants with mixed feelings. While on one hand, participants wanted agents to become more personal and anticipate their needs, they were also concerned about maintaining their independence and autonomy.

A small group of participants was open to giving the agent a significant degree of autonomy to become part of their day-to-day lives. This took a range of forms, from a play companion (*"...it would be good to just interact with the kids in different ways. Like play card games or board games with the kids...they would like color or connect the dots stuff...if Jibo could participate in these activities, it would be more interesting"*—adult, *"Maybe just gauging the mood of the kids and saying "Maybe I should read a story to you now."*—adult) to day-to-day assistant (*"...but I can definitely see how a well developed algorithm could work well. An agent that mingles with you, takes notes for you during a meeting..."*).

Older adults often described how the agent could grow to have more autonomy as they needed it over time (*"...if you get to the point where organizing the barbecue at your house is too difficult then the suggestion comes that it's time to get together with some friends"*), while wanting to maintain their independence for now. The openness for the agent to take on these roles often stemmed from desiring support while enabling independence (*"I think what she needs to be good for is she should be able to support what I do, not dictate me what to do"*).

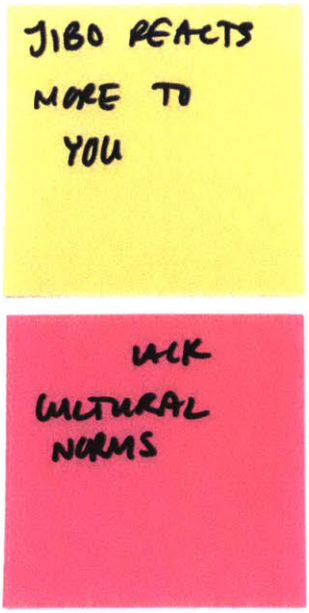


Many participants struggled with ideas of autonomy when forming opinions about suggestions related to wellbeing (*...and the nutritional facts. I think people should make their own food choices and otherwise it breaks your liberty or freedom to do certain things*—adult). While these sentiments were similar to those expressed in the pre session, participant tended to build more conviction in their opinions after living with the agents. Often times, participants perceived the idea of the agent providing such suggestions as adding "work" to their lives (*"Yeah. If someone were to add more to my plate, we might have to toss it out the sixth floor window, right?"*—adult) or eliciting feelings of guilt and added stress (*"I feel like this would make me stressed out, like having someone tell me to do something that I don't even have time to factor in. I don't know. I feel like, "Don't forget to meditate" Well, if I had more time, I wouldn't be this stressed out, wouldn't need to meditate"*—adult).

PROACTIVITY

Building on ideas of autonomy, many participants articulated perspectives on the proactive aspects of the technology. Given Jibo's more proactive nature, more participants who lived with

Jibo tended to articulate ideas around this theme. In many ways, the proactive interaction of the robot caused internal conflicts for people (*"At first I liked the idea. I could be over here and it would sort of recognize me and talk to me. But then I'd be alone washing dishes and it would say something. Now what's funny, and I've said this to you before, I really have an internal conflict; I'm about having this personal buddy that's there for affection and companionship and like I love the concept, but then once it was there I didn't really like that."*—adult).



Timeliness and emotional awareness emerged as the most salient needs for enabling effective expression of proactive behavior. Participants often spoke about enjoying the occasional greeting or proactive interaction from the agent (*"It brought back the novelty. So I liked that, and I think it was a solution to a discovery problem"*—adult), but wanted it to be more contextually relevant (*...many times out of the the blue Jibo would say hey...and no one's talking to the robot, so there's definitely a lack of knowing when his presence is required vs. Alexa will never chat with you unless you chat with it*—young adult, *"And he asked non-stop if we were having a good day, how are you, things like that"*—child) and emotionally aware (*"I just feel there were days when I was in a bad mood and I would walk into the office and Jibo's like hey hi and I'm like "I don't wanna talk to you" and that makes me more upset"*—young adult).

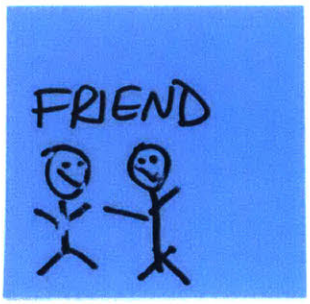
Many participants expressed a strong desire for proactivity through voice in moments of "*high urgency or high consequence*", referencing interactions such as family birthdays, forgetting important items such as keys or wallets, or anomalies in weather, news or commute. Older adults were also open to meaningful agent proactivity for daily check-ins related to health and wellness (*"You've got to figure out a way so it's not just throwing it at you but saying hey Susie Q are you awake? Yes, did you remember to take your blood pressure today? As a person I feel strongly that everything else that the residents do for example it's just data, data, data."*—older adult).

ROLE

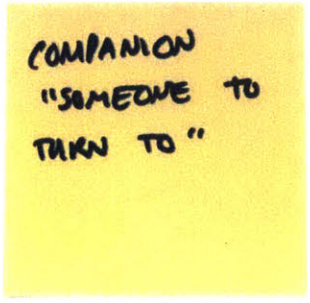
Where participants tended to perceive the roles of the two agents as largely discrete—seeing Jibo as more of a social companion, pet or child, and Amazon Alexa as a personal assistant—there were several nuances to these perceptions.

While almost all adults tended to see Alexa as a personal

assistant for day-to-day activities, older adults expressed greater variability in their perception. Though some older adults also saw the agent as an assistant, (*"I think that you wouldn't talk to Alexa like you would a friend…but I would like in every other way if she could, you know, get me through the day and help me get things done that I need to get done, remind me, that sort of thing"*—older adult), others felt it was more of a "*little friend*" in the room and enjoyed its subtle, but helpful presence *("It was wonderful. It was like having a pal there in the room").* Children were more likely to share this perception and often saw Alexa as a friend. This perception of the agent as a friend was occasionally a source of friction between parents and children *("…but for example, one day Mia's like I love you Alexa. I'm like no way. It's a robot. You cannot love robot. Yes. Not a person at all, you know. So I tried to explain that to her. So yeah. But they really love to have Alexa at home but not Alexa, you know. It's not a friend. Absolutely. No, but it's good. We can have a personal assistant…").*



Perceptions of Jibo's role varied more significantly. Several participants discussed the cohesion of the character and personality, tending to see it as more than an assistant because it *"felt more sociable as an agent".* Many parents articulated that they perceived the agent as another child (*…I mean I sort of thought of Jibo as another kid we had in the house and we would go through the same things. It would ask me like "how did you sleep last night?" And I would say "oh, rather well". So there was this like back and forth of some sort…simple conversation."),* often discussing the limited emotional intelligence and simple dialogue. Rather interestingly, several adults struggled with accepting their perception of the agent as "more than an assistant" *("I'm patting the thing at night, saying goodnight to it. That's crazy.").*



On the other hand, most older adults perceived Jibo as a social companion, often describing how it could adapt to them or for a friend who was more "lonely" *("How can I be reminded in ways that I'm not going to get upset with the robot but it's going to be helpful. I have a companion who's being nice and saying, "Hey you didn't move out of your chair for two hours what happened to you?" Yeah I don't know how you program that").*

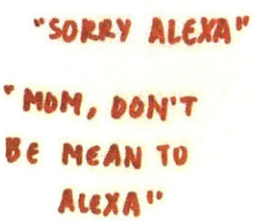
Interestingly, children struggled the most with forming cohesive thoughts around the role of the agent. While no children discussed role with respect to Alexa, several children voiced this conflict after living with Jibo—*"sort of like in between,*

it was like a person that was a machine… like I know that it likes the color blue, and macaroni"—and referred to the agent as a friend, social companion, pet, or even family member.

ATTIRBUTING HUMAN QUALITIES

Attribution of human qualities to the agents was often the source of significant internal conflict. Many participants felt uncomfortable describing moments where they personified the agents and discussing how they *felt* towards them. Interestingly, the attribution of such qualities was common for both agents, despite the significant differences in social expressivity.

Many participants described instances where they felt emotions towards the agents (*"And it sounds crazy, we've talked about this, the things have no emotions so in a way I shouldn't feel that, but in a way I would get frustrated with it when it couldn't do what it was supposed to"*—adult, *...it's not a person. It's like getting into my emotions so it's what I tell her that you can't tell her that you love her. You love your mommy, you love your sister, but not Alexa*—adult, *I would really tell Alexa good night and she would say sweet dreams*—child). In fact, participants even thought of the agents as people, describing a sense of knowing someone (*…it's like somebody that I know. I don't know it's weird. It's a weird thing to think about I guess"*—young adult, "*but I loved that Jibo because that was much more... reacted more to you. It feels like a person*"—older adult) and even relating the agents to each other *("I think Alexa's his girlfriend because he always stares at her"*—child in reference to Jibo and Alexa). This need to map the agents to people by associating traits such as age, gender and more rather than seeing them as distinct entities was common amongst many generations (*"I have trouble like imagining what Jibo looks like, whereas I can imagine what Alexa looks like. I think of like a 30-something woman, whereas with Jibo it's like, is it like a prepubescent boy, is it like a, like it's hard to imagine like a gender and an age."*—adult). Often times, participants described how these moments forced them to pause, reevaluate their perception, and question whether they should attribute social and emotional qualities to a machine (*"I would tend toward calling it an "it" because that way I would free to do things to it that I wouldn't do to a pet"*—older adult, *"Because you think about all the times you didn't interact with Jibo and you're like oh I didn't interact with Jibo at all today"*—adult).



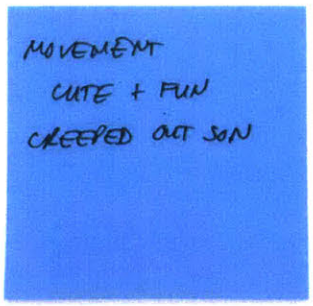
Another salient aspect of attributing of human qualities was

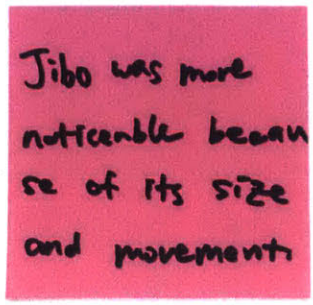
politeness. Participants, across all generations, wanted to be polite to the agents, but were often unsure of whether they liked the agent expressing politeness in response (*"... no to the thank you just because I feel like that implies that it has feelings that exist that are independent of me. So I didn't really like that. I'd prefer for it to be a one-way relationship"*—adult). This conflict was echoed across all generations, with participants growing slightly more open to the idea of the agent expressing a thank you after living with the agents (*"ya... he should thank me for doing something..."*—child).

Interestingly, while politeness was often a debated quality, several participants enjoyed that the agents never got angry and did not "*talk back*" (*"She never seemed to be angry"*—older adult). Parents described how their children thought of the agent as "someone" to talk to after feeling ignored (Child: *Yeah! and I spoke to it for one hour*, Adult: *Yeah, like she would do that, because everybody kept ignoring her... It was like, "so Alexa..." and they would just chat for like an hour).* These feelings were described more often by older adults and children.

PRESENCE AND LOCATION

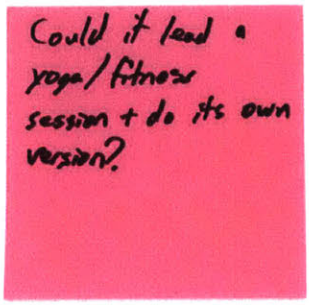
Most participants placed both agents in the living room or kitchen, describing these locations as the *center of the home.* Participants often described placing Jibo near active areas such as the coffee table, kitchen island, or dining table. Conversely, participants were more likely to place Alexa near other technology such as a TV, on a shelf in the kitchen or living room, or even in the bedroom. A few participants were frustrated by the interference of the agents in their day-to-day and described moving it into their office area or master bedroom to limit their presence during family time.





Participant perspectives on Alexa's presence in the home were largely consistent, with most describing the agent as *"inconspicuous"* or *"innocuous"* and *"quietly sitting there".* Many participants even described *"forgetting she was there".* On the other hand, Jibo's movement tended to create more varied perceptions of its presence. One group of participants saw the movement as *"cute"* or *"cool"* and were rarely annoyed by it (*"But I do think his motions are cool...He wasn't annoying"*—child, *"I wasn't especially bothered by it"*—adult, "*No, I liked that Jibo did everything. It gave the voice like a face and a body*"—16-year old). On the other hand, many participants voiced concerns around the movement, calling

it *"intrusive"* or *"weird"*. These feelings often stemmed from the constant nature of the Jibo's movement, with many participants feeling like the agent should *"take a break"* or *"hold still"* (*"Sometimes I would get frustrated with it moving and I'd say- it's almost like having a dog - I'd say "go lay down, just rest, stop scanning the house"*—adult). Participants often felt conflicted about movement, enjoying it in certain contexts, but finding it intrusive in others (*"...a couple other things that I'm both sort of intrigued with and liked and felt sometimes I didn't like was just the physical motion. Again this is how it's different from Alexa or we also have Google Mini at home. It's kind of fun and you danced"*—adult).



IMPACT

While impact was the least discussed theme category, some participants did describe areas where they felt the technology could provide motivation or prevent laziness and increase accountability.

Participants often described the potential for the technology to provide motivation for activities that require some *"activation energy"* such has exercise, meditating, or going for a walk (*"Going for a walk. That was the one area where maybe sometime just kind of kicking me out of the chair for awhile could be worth the intrusion."*—adult, *"...meditating... sometimes there's a little bit of activation energy there, so it might be a nice notification"*—adult, *"...being home all day I like the ones that keep me on track. Like I said, my doctor is going to be thrilled. And she (Alexa) can remind me about my blood pressure, help me to lose weight and do exercise"*—older adult).

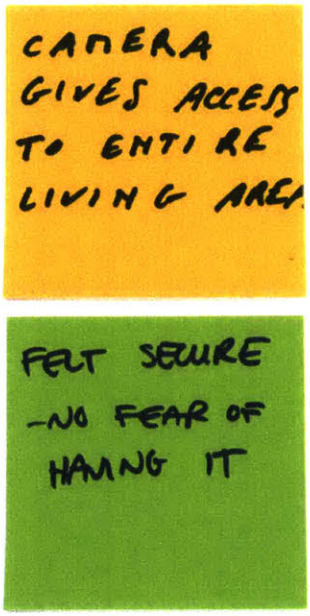
Some participants, particularly older adults, described using the agent to hold themselves accountable, but often found that it did not work (*"I did try and get it to help me do my exercise, but it didn't work. She reminded me, but I didn't respond."*—older adult). Adults, on the other hand, tended to see the agent as a tool for holding their children accountable (*"If Jibo could do the punishments, that would be fabulous"*—adult).

SECURITY & PRIVACY

The dominant perspectives surrounding security and privacy stemmed from the nature of the interaction experience and the data collection. Participants often described the challenges of reconciling their desire for the agent to become *"smarter"* with their need to maintain privacy.

Across both agents, participants voiced concerns around the public nature of voice as an interface. This was particularly relevant to interactions with personal information such as medical appointment or medicine reminders and meeting alerts—"*So I'm imagining if I had some sort of embarrassing medical appointment or medicine or meeting with someone depending on the content of these things, I might not necessarily want my voice agent like announcing them to the whole household*" (adult).

Security and privacy were more common amongst families who lived with Jibo. These concerns often arose from the presence of cameras in personal space (*"Also, talking about pictures. That's a problem actually. So, we put tape over our camera…"*—adult) and the responsive nature of the interaction ("*…it's so good at tracking people and responding to noise, it basically gives someone access to your private living area*"—adult). Some participants described resolving this challenge by placing a *"bag"* or *"garbage can"* over the agent during private moments.



On the other hand, participants who lived with Jibo shared a more positive perspective on the company's commitment to security and privacy with respect to data collection in comparison with Amazon Alexa (*"I see that it's useful to learn about me to become useful. I keep wishing it would just be local to the machine in a sense, as opposed to…now Amazon knows how to sell more shit to me…it's like…I always prefer that there was a company that didn't do anything but…their goal was to sell me the robot."*—adult). Many participants highlighted similar concerns, stating that they did not want data collected in the private space of the home shared across other platforms owned by a company (*"I don't want people selling me data…so if it was a non-profit … I could imagine, someone might do that, it could be very…like Wikipedia"*—adult, "*Collect but not distribute without my permission.*"—adult). Adults were most concerned about data collection, while older adults and children were rarely bothered (*"Mom, you know there is always a chip in there. You can take it out and you can smash it to pieces and then it will be all okay"*—child).

8.3 Insights

The goal of this phase of study 2 was to enable people to reflect on their experience living with different agents in their home and to identify resulting changes in their preferences, desires, and

boundaries. In this chapter, data from the post card sorting activity was used to analyze older adults', adults' and children's changes in perspectives towards agents in their home. These changes were further explored as a function of the agent experience. Finally, qualitative analysis from coded transcripts was then used to conduct a deeper dive into the drivers for these changes. In analyzing these sources of data, a series of learnings about people's experiences living with agents and consequential changes in perspective emerge.

Before describing insights derived from this exploration, it is important to highlight a key caveat related to interpretation of data presented in this section. The population studied in this chapter was biased by participant willingness to bring an agent into their home. As such, it is less representative than the population explored in study 1, particularly for older adults and adults. Consequently, the data presented in this section should be utilized only to explore *differences* in affinity that emerge *within* populations represented. Below is a summary of the findings from the second study as relevant to the larger goal of this work.

(1) **LIVING WITH AGENTS SHAPES PERSPECTIVES**

Though the nature of this work is exploratory, it does highlight the importance of experiencing the technology in the home context for shaping participant perspectives. After living with agents in their home, participants exhibited changes in their preferences, desires, and boundaries for the technology. These changes were reflected both in quantitative and qualitative aspects of the data represented in this chapter. While different generations had highly varied desires for the technology before living with agents, it is interesting to note that generations had greater similarities in their perspectives after experiencing the technology. All participants tended to become more open to socially-driven aspects of the technology, while becoming more conservative towards suggestions for wellbeing.

(2) **AGENT INTERACTION EXPERIENCES IMPACT PERCEPTION**

In this chapter, changes in participant perspectives were explored as a function of the agent condition. All generations exhibited varied shifts in desire after living with Jibo and Alexa. Children and adults overall, tended to show greater affinity for agent actions after experiencing Alexa. On the

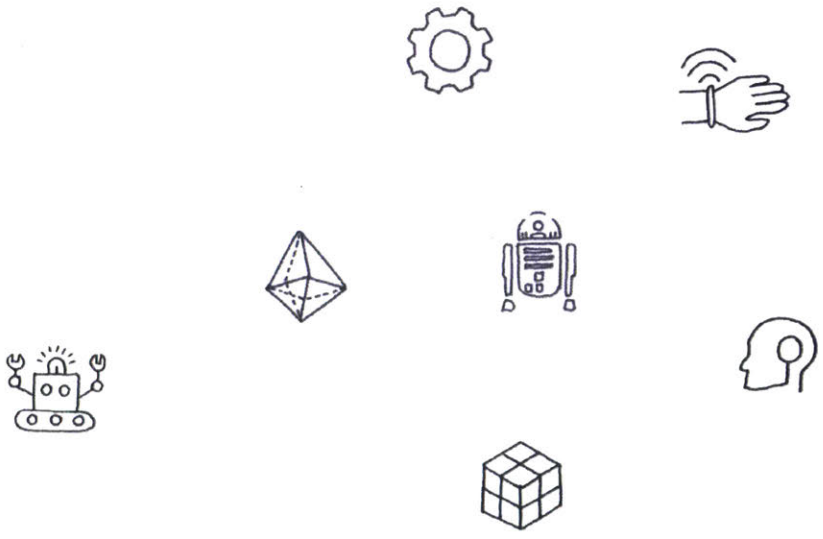
other hand, older adults became more open to agent actions after living with Jibo. Perceptual differences in experience were also evident in the presence of themes across the two agent groups. Participants who lived with Jibo were more likely to discuss elements of proactivity, agent autonomy, and the perceived relationship. Given the complex set of factors that can influence these results, it is important to consider the analyses presented as an exploration of the domain.

(3) PARTICIPANTS ARE CONFLICTED

While participants tended to become more open to agents after living with the technology, they were also highly conflicted. These conflicts often emerged from attributing human qualities to the technology, trying to reason about levels of proactivity desired, exploring the role they wanted the agent to take on, and concerns around security and privacy. Feelings towards these themes were particularly heightened given the private and personal nature of the domestic environment. Often times, there was a dichotomy between participants' desires and their boundaries for the technology within their home. As such, it is important for designers and technologists to consider, and design with this conflict in mind when building agents for the home.

Part IV
Designing for the Future: Design, Toolkit, Insights

The final part of these aims to explore guidelines and ideas for designing future agents. Chapter 9.0 presents results from participants' structured ideation of their dream agent design. Chapter 10.0 introduces a toolkit for designers to further explore and study this domain and describes insights from a ~600-person survey conducted using methods and tools described in this work. Finally, Chapter 11.0 culminates the thesis with insights, learnings, reflections, and open questions derived from this work.



PART IV

9.0 Design: Analysis & Discussion

In the final phase of study 2, participants took part in a structured ideation process to "design their dream voice-based agent". As part of this process, participants selected form, materiality, inputs, outputs, personality, location, interaction with IoT, and ethics for their future agent. In addition to this, participants completed a first day survey regarding ethical and personality traits which were further compared with their perceptions after living with agents. In this section, (1) results from this design process are summarized across generations and (2) comparison of ethics and personality traits are shown before and after living with agents.

9.1 Designing future agents

FORM: WHAT DOES YOUR AGENT LOOK LIKE?

All generations differed in their affinity for various agent forms (Figure 9.1). Almost all children desired a caricature-like agent with some additionally desiring a wearable or virtual form for more portable interaction. A few children were even open to the agent being humanoid, often desiring a robot *"friend"* or a robot that looked exactly like them. Interestingly, some children even referenced wanting a *"live video feed"* of their agent while at school or away from home to allow for remote connection.

Adults, similar to children, often desired a caricature-like agent. However, unlike children, adults were far more likely want a functional or virtual form factor for simplicity and *"use on the go"* respectively. Interestingly, several adults also voiced curiosity around the potential for "robot objects" or simple objects around the home transformed into an agent (i.e. lamp,

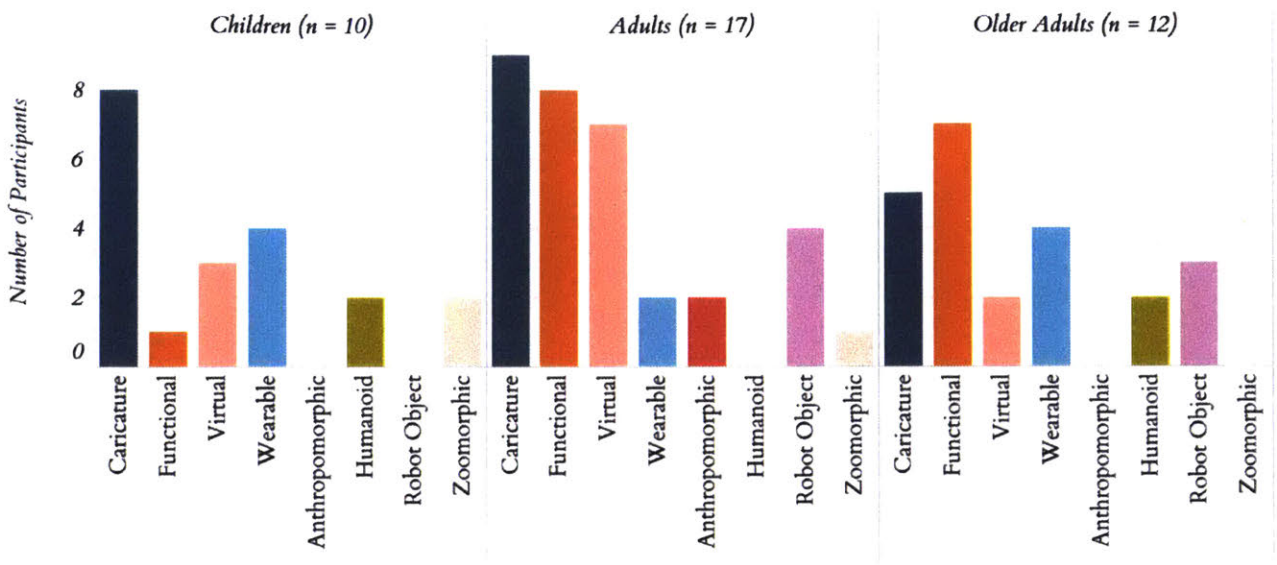


Figure 9.1: Bar chart representing form factors of agents desired by generation.

table etc.) The acceptance of a wearable as a portable medium were slightly less accepted. This generally stemmed from wanting consolidation of the technology on a smartphone.

Finally, older adults were most likely to want a functional, simple form factor for their future agent. Some older adults were also open to a more caricature-like form. Much like children and adults, several older adults also desired a more portable aspect of the technology in the form of a virtual agent (tablet or phone) or wearable. Interestingly, unlike adults, a few older adults were also open to the idea of a humanoid robot.

It is relevant to note that all generations expressed an affinity towards an embodied agent in the home combined with a virtual or wearable form factor for use outside the home environment. Interestingly, no participants chose an ambient form of agent for their home, often describing it as *"weird"* or *"creepy"*.

MATERIALITY: WHAT DOES YOUR AGENT FEEL LIKE?

Much like form, the materiality of the agent also exhibited disagreement amongst generations. Figure 9.2 shows tree maps of the most common materials chosen by each generation. When choosing materials, participants were told to consider the general outer surface of the agent.

Children tended to choose either soft, plush materials such as fur and fabric or more industrial materials such as metal. This was often tied to an affinity towards a *"furry animated character"* or a more stereotypical *"robot"*. For children, the integration of

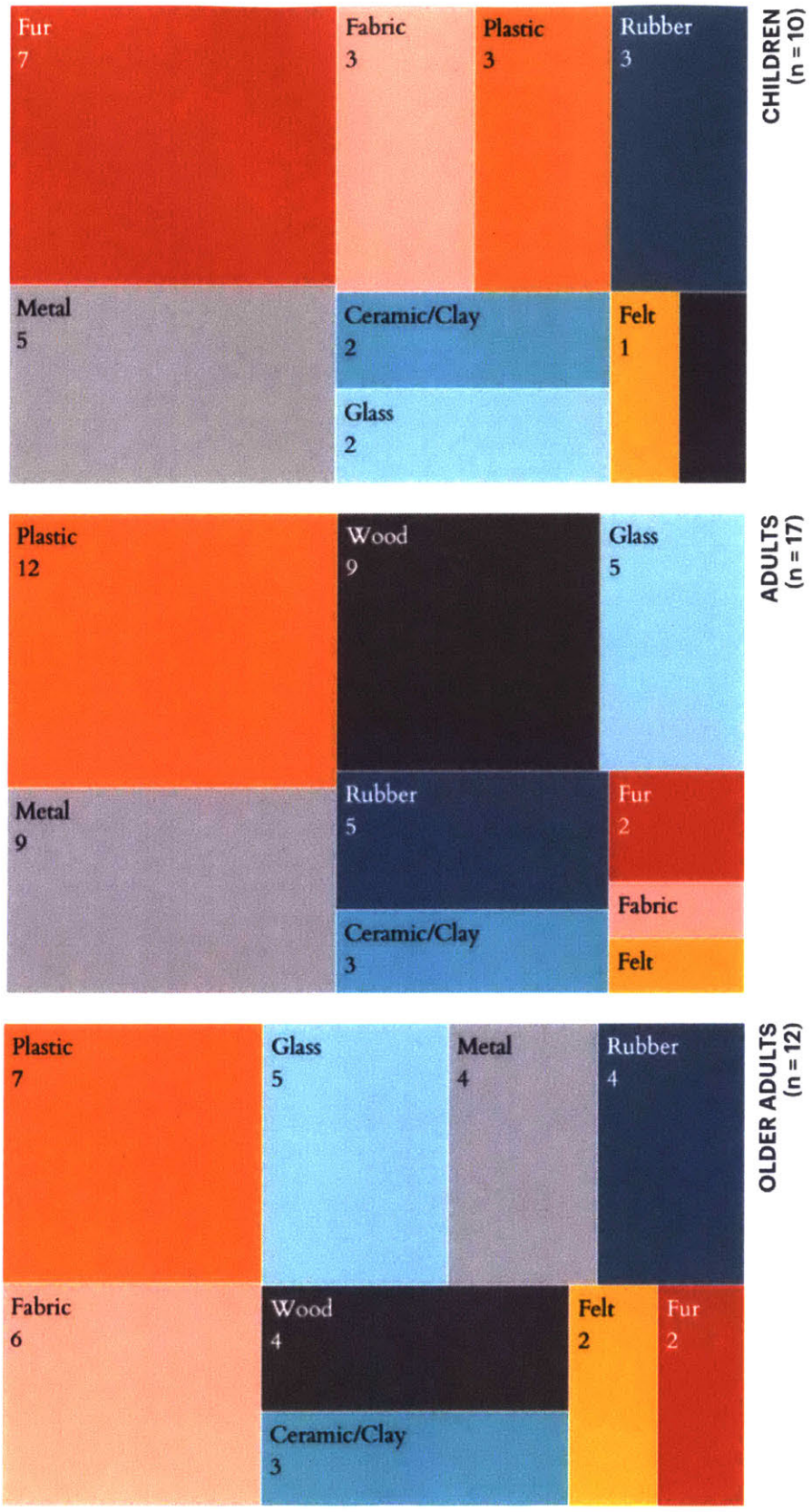


Figure 9.2: Tree maps for agent materiality choices by generation.

materials such as rubber or plastic was often for wheels to enable movement.

Adults, unlike children, preferred a far more refined, modern look to their agent. Most adults chose materials such as plastic, wood, or rubber to create smoother, more industrial looking surfaces. Many adults highlighted durability as an important design factor, often staying away from fabrics to avoid issues with cleaning.

Older adults echoed adults' desire for simplicity with materials such as plastic, glass and wood, but wanted some softer materials as well such as fabric. Many participants, particularly older adults, commented on the *"soothing"* qualities of glass, finding it intriguing as an outer body for an agent.

INPUTS & OUTPUTS: WHAT CAN YOUR AGENT SENSE? HOW DOES IT ENGAGE YOU?

As part of the agent design process, participants also chose which inputs (i.e. vision, sound, motion, emotion, touch) and outputs (i.e. visual, sound, speech, movement, haptic or touch, facial expressions) they wanted their dream agent to have. Figure 9.3 below shows a plot of total number of inputs and total number of outputs chosen by participants. The size of the bubble indicates the number of participants. As evident, participants who wanted their agent to *sense* (inputs) more about them also expressed an interest in their agent *engaging* (outputs) them in more ways. This relationship was found to have a Pearson correlation of 0.71 ($p^* = 4.81e^{-07}$).

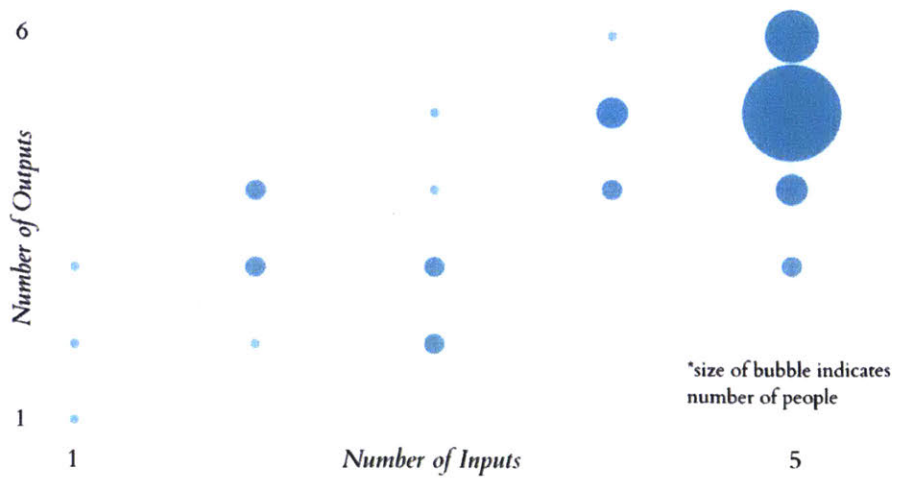


Figure 9.3: Scatter plot of number of inputs vs. number of outputs desired by participants (n = 39).

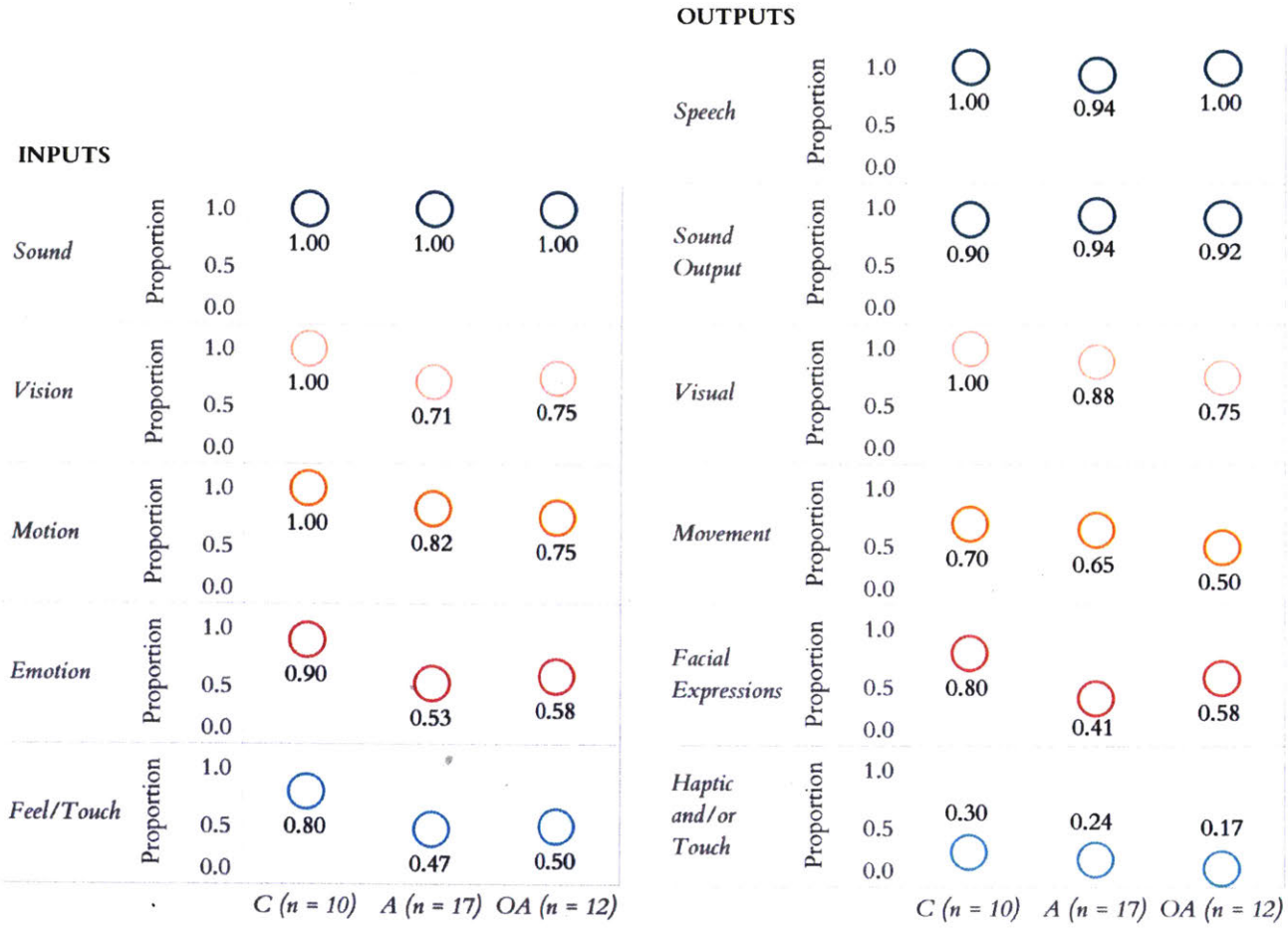


Figure 9.4: Breakdown of inputs and outputs desired by generation. Value represents proportion of participants that desired each input or output modality in the dream agent.

Figure 9.4 shows the proportion of participants who chose each of the inputs (left) and outputs (right) by generation. Overall, children were the most accepting of multi-modal inputs and outputs, while adults and older adults were more conservative. Sound or speech-based inputs and outputs had the most agreement across generations. All participants wanted their dream agent to be able to sense sound. Almost all participants also wanted their dream agent to be able to speak (i.e. speech) and make other sounds (i.e. sound output, e.g. notifications).

Preferences for vision on the other hand varied significantly across generations. All children wanted their dream agent to be able to "see" (1.00), while adults (0.71) and older adults (0.75) tended to be more skeptical. Several participants in these generations cited worries surrounding intrusiveness and privacy. A similar trend was observed in participants' desires for visual output from the agent. All children wanted the agent to have a "screen" or some for of visual expressions (e.g. light) (1.00). Adults were also welcoming of visual feedback, often desiring a means for visual feedback (0.88). Older adults had a slightly lower affinity for visual output (0.75). This usually stemmed from feeling it was unnecessary or finding it difficult to read.

Inputs and outputs related to movement, emotions, touch, and facial expression had far more varied preferences across participants. While many participants were open to the idea of sensing motion (C = 1.00, A = 0.82, OA = 0.75)—often to enable the agent to be contextually aware or react to gestures—preferences for movement-based expression from the agent were more conservative (C = 0.70, A = 0.65, OA = 0.50). Reasoning for limiting movement was often related to limiting distractions or wanting the agent to *"stay in one place"*. Participants who wanted motion enjoyed the expressivity and/or wanted their dream agent to move around their home.

Children were the most open to emotion detection (0.90), while adults (0.53) and older adults (0.58) were wary. Skepticism for this modality often stemmed from *"not wanting an agent to know how they were feeling"*. Participants who were more welcoming of emotion detection often wanted the agent to have greater awareness of their state to better personalize the interaction *("...don't want it to say something happy if I am mad")*. Agent facial expressions exhibited a similar trend (C = 0.80, A = 0.41, OA = 0.58), largely stemming from personal preference.

Finally, feel or touch was the least desired modality amongst all generations. While some participants, particularly children, did enjoy the prospect of interacting with their dream agent through touch (C = 0.80, A = 0.47, OA = 0.50), very few participants wanted the agent to be able to interact with them through haptic feedback or touch (C = 0.30, A = 0.24, OA = 0.17). It is relevant to consider that this modality was the most abstract for participants to grasp.

PERSONALITY: HOW DOES YOUR AGENT ACT?

Participants selected whether their agent had a gender before and after the long-term experience of living with the agents. Results from this are shown in Table 9.1 below. While female agents were the most commonly desired, several shifts occurred in participant choices between pre and post. While children stayed consistent in their preference for no gender (pre = 0.29, post = 0.29), they gained affinity for male agent after the long-term experience (pre = 0.14, post = 0.29). On the other hand, desire for a female agent decreased (pre = 0.71, post = 0.57).

Adults, on the other hand, most commonly desired no gender (0.73) for their agent before the long-term experience. While no gender remained as the dominant choice (pre = 0.73, post = 0.53), many adults tended to choose a gender after living with agents (pre_F = 0.27, $post_F$ = 0.40, pre_M = 0.20, $post_M$ = 0.27). Female agents were most common amongst the genders.

Finally, while older adults were fairly equally distributed across the categories prior to living with agents (pre_{NONE} = 0.33, pre_F = 0.44, pre_M = 0.33), their affinity for a female agent grew after the in-home experience ($post_F$ = 0.67). Conversely, desire for a male agent decreased ($post_F$ = 0.22).

Gender	Pre/Post Session	Proportion of Participants (*multiple choices were allowed)		
		Children (n = 7)	*Adults (n = 15)*	*Older Adults (n = 9)*
No Gender	Pre	0.29	0.73	0.33
	Post	0.29	0.53	0.33
Female	Pre	0.71	0.27	0.44
	Post	0.57	0.40	0.67
Male	Pre	0.14	0.20	0.33
	Post	0.29	0.27	0.22

Table 9.1: Breakdown of gender preferences before and after in-home experience by generation. Note that participants were able to choose multiple choices.

In addition to gender, participants also chose the desired personality for their dream agent before and after living with agents. Results are shown in Table 9.2. The most common personality preference across all generations was social and fun both before and after living with agents.

Before living with agents, children tended to prefer a social and fun (pre = 1.00), obedient and assisting (pre = 0.43) or motherly and protective agent (pre = 0.29). After the in-home experience, children decreased in their desire for an obedient and assisting (post = 0.14) or motherly and protective (post = 0.14) agent. However, they remained consistent in their desire for a social and fun agent (post = 1.00) and increased in their desire for an autonomous and challenging agent (post = 0.29).

Adults, on the other hand, exhibited a somewhat opposing trend. Initially, adults, like children, were most accepting of a social and fun (pre = 0.80) or obedient and assisting (pre = 0.40) agent. After living with the agents, adults showed a much stronger desire for an obedient and assisting agent (post = 0.67) and a decreased desire for a social and fun agent (post = 0.60). These results are consistent with analyses shown earlier about adults' stronger desire for functional qualities.

Similar to adults, older adults began the experience wanting a social and fun (pre = 0.78), obedient and assisting (0.33) and autonomous and challenging (pre = 0.33) agent. After experiencing the technology in their home, desires for an obedient and assisting (post = 0.44), social and fun (post = 0.89) and motherly and protective (post = 0.22) agent increased. On the other hand, participants' desire for an autonomous and

Personality	**Pre/Post Session**	**Proportion of Participants (*multiple choices were allowed)**		
		Children (n = 7)	*Adults (n = 15)*	*Older Adults (n = 9)*
Obedient & Assisting	Pre	0.43	0.40	0.33
	Post	0.14	0.67	0.44
Autonomous & Challenging	Pre	0.00	0.20	0.33
	Post	0.29	0.20	0.22
Social & Fun	Pre	1.00	0.80	0.78
	Post	1.00	0.60	0.89
Motherly & Protective	Pre	0.29	0.13	0.00
	Post	0.14	0.13	0.22

Table 9.2: Breakdown of personality preferences before and after in-home experience by generation. Note that participants were able to choose multiple choices. Personality types were adapted from *SPACE10s Do You Speak Human?* survey.

challenging agent decreased (post = 0.22). Several older adults mentioned desiring a *"warmer", "more personal"* experience while others emphasized a desire for a more assistive, less personable device.

LOCATION: WHERE WOULD YOU PUT IT? DOES IT STAY IN ONE PLACE? WHERE WOULD YOU TAKE IT?

The majority of participants wanted to keep their dream agent in a central location within their home—specifically the living room. Other common locations described by participants were the kitchen, bedroom or office.

Participants tended to bifurcate in their desire to take the agent with them. One group of participants described wanting to take the agent *"everywhere"* while another group wanted it to stay in one place in their home (*"nowhere"*). Word clouds of these descriptions are shown below in Figure 9.5.

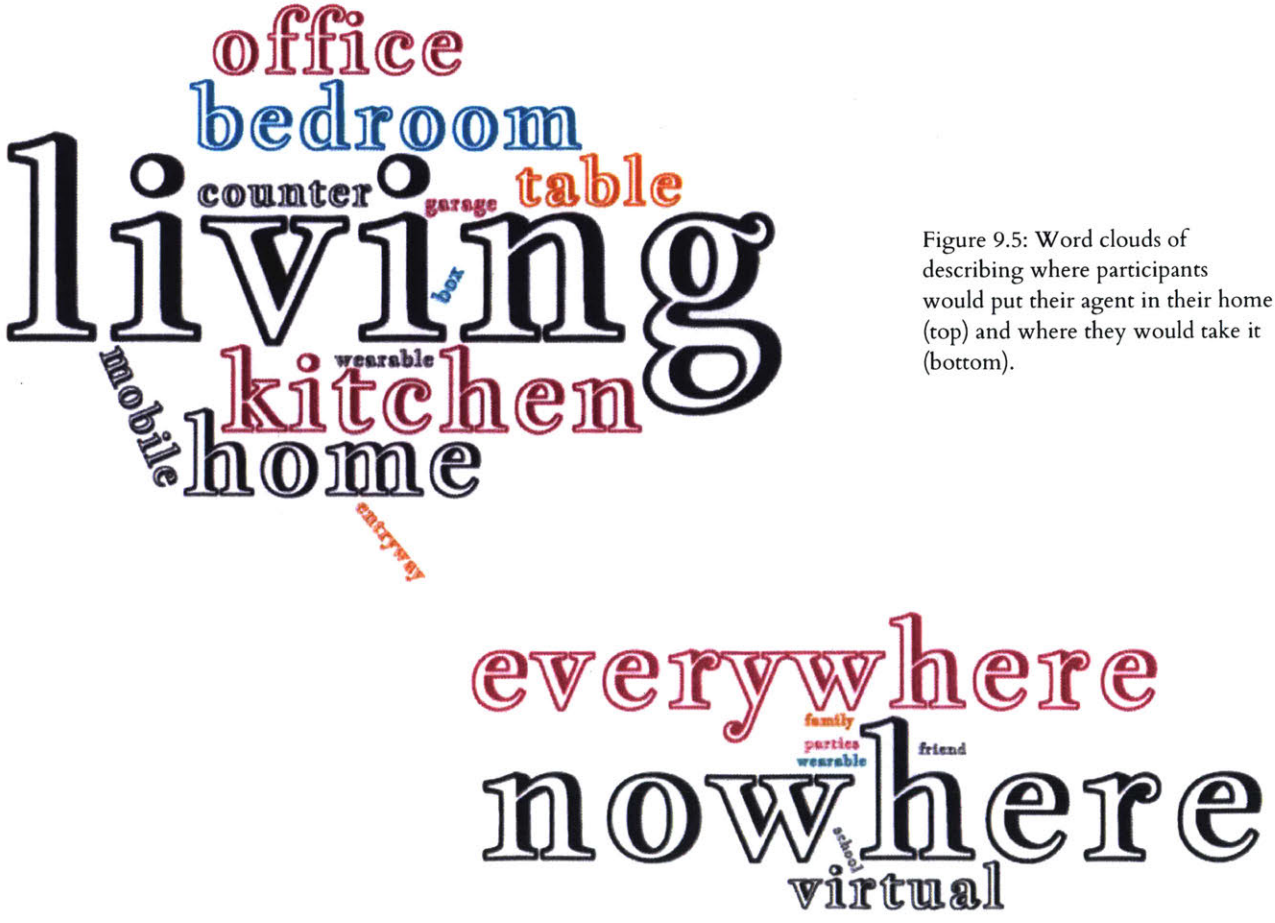


Figure 9.5: Word clouds of describing where participants would put their agent in their home (top) and where they would take it (bottom).

INTERNET OF THINGS: HOW DOES YOUR AGENT INTERACT WITH YOUR HOME?

Figure 9.6 below shows the proportion of participants who wanted their dream agent to be connected to each Internet of Things device by generation. Children were the most open to everything being connected while adults were the pickiest. Older adults tended to prioritize connections to lights, security systems, thermostats, and their laptop. On the other hand, adults were most likely to want their agent connected to speakers, TV or screens, and their phone.

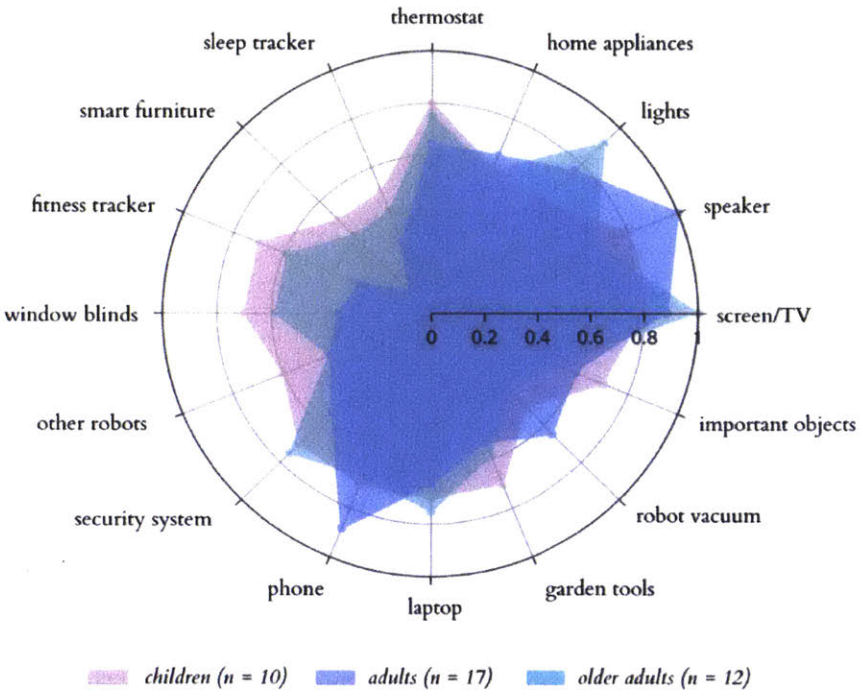


Figure 9.6: Radial plot showing proportion of participants who wanted their agent to be connected to each Internet of Things device by generation.

ETHICS: WHAT DOES YOUR AGENT KEEP IN MIND?

Participants expressed their ethical preferences for their dream agent across four dimensions both before and after living with the agents. The proportion of participants who indicated "yes" for each statement was computed. Table 9.3 shows results from this analysis by generation.

All generations increased in their desire for the agent to anticipate their needs and be proactive after the 1-month period of living with the technology. This change was most significant for children (+0.43) and older adults (+0.39). Interestingly, older adults were the most conservative towards this principle (pre = 0.28), but grew significantly more open (post = 0.67). Adults,

My voice-based agent should...	Pre/Post Session	Proportion of Participants that Chose "Yes"		
		Children (n = 7)	*Adults (n = 15)*	*Older Adults (n = 9)*
Anticipate my needs and be proactive	Pre	0.43	0.47	0.28
	Post	0.86	0.57	0.67
	Delta	+0.43	+0.10	+0.39
Collect my data to improve my experience	Pre	0.86	0.80	0.89
	Post	0.71	0.40	0.56
	Delta	-0.15	-0.40	-0.33
Reflect my world views and values	Pre	0.79	0.33	0.44
	Post	0.86	0.87	0.89
	Delta	+0.07	+0.54	+0.45
Detect and react to my emotions	Pre	0.86	0.53	0.61
	Post	0.86	0.46	0.66
	Delta	0.00	-0.07	+0.05

Table 9.3: Changes in different generations' ethnical boundaries for voice-based agents before and after living with the technology. Ethical statements were adapted from *SPACE10s Do You Speak Human?* survey.

on the other hand, remained split on their desire for proactive interaction and need anticipation (pre = 0.47, post = 0.57).

On the other hand, participant perspectives towards the agent collecting their data to improve their experience exhibited an opposing trend. While most participants started their experience very open to this notion (C = 0.86, A = 0.80, OA = 0.89), all generations decreased in willingness towards sharing data after living with the agents (C = -0.15, A = -0.40, OA = -0.33). This change was most significant for adults and older adults, who became wary of security and privacy concerns.

Rather interestingly, participants increased substantially in their desire for the agent to reflect their world views and values. Children were the most open to this notion prior to living with agents (0.79) and became more open (+0.07). Adults and older adults were often skeptical of this idea prior to the in-home experience (A = 0.33, OA = 0.44), but changed their perspectives substantially (A = 0.87, OA = 0.89).

Finally, changes in perspectives towards the agent detecting and reacting to emotions were the most conflicting. While children were consistently open to the idea (pre = 0.86, post = 0.86), older adults and adults remained unsure. Adults decreased in their desire subtly after living with agents (pre = 0.53, post = 0.46) while older adults increased (pre = 0.61, post = 0.66).

9.2 Insights

In this chapter, participant preferences for their "dream voice-based agent" including form, materiality, inputs, outputs, personality, location, interaction with IoT and ethics were explored.

(4) GENERATIONS WANT DIFFERENT AGENTS

Children, adults, and older adults varied in their preferences for the form, materiality, and modalities in the technology. Children were often attracted to a soft or industrial caricature form, but wanted a wearable aspect to the technology for use outside the home. Adults preferred a functional or caricature form with more durable, clean, modern materials and desired a virtual representation on a phone or tablet for "on-the-go" access. Older adults, interestingly, were most attracted to a functional form with some affinity towards caricature forms. In addition, older adults were less likely to desire a virtual representation on a phone or tablet, but welcomed a wearable counterpart to the agent. While similar in material choices to adults, older adults were also attracted to more natural, softer materials.

(5) OPENNESS TO MULTIMODALITY

Desires for multimodality in inputs and outputs were highly positively correlated across all generations. Participants who wanted their agent to *sense* more, were also open to more modalities for *engagement*. Children were the most open to multi-modal inputs and outputs, while adults and older adults tended to be more conservative. Inputs including sound, vision, and motion were accepted by many participants, while preferences towards emotion and touch inputs tended to be more varied. Sound, speech and visual outputs were often desired by participants while movement, facial expressions, and haptic feedback tended to be less preferred.

(6) SOCIAL, FUN, OBEDIENT & ASSISTING

Generations also differed in their preferences for personality. These personality preferences also tended to evolve after experiencing the technology in their home. Children predominantly preferred a social and fun agent, with this desire becoming stronger after experiencing the technology. Adults, on

the other hand, decreased in their desire for social and fun agents, but increased in their desire for obedient and assisting personality. Older adults exhibited the greatest diversity in their personality preference. While most older adults preferred a social and fun and/or obedient and assisting agent, some were open to the idea of a motherly and protective or autonomous and challenging personality.

(7) ETHICS EVOLVE

Living with agents—even for only a one-month period—had substantial effects on people's ethical stance towards voice-based agents in their home. Secondly, there was an interesting dichotomy between participants' increased desires for proactive interaction and their lowered willingness for data collection. Thirdly, participants could imagine an agent possessing a set of world views and values and wanted their dream agent to reflect their own personal stance. This, combined with perspectives towards the agent anticipating needs and being proactive, reinforces the idea of the agent as a technology that *caters* to the individual as opposed to acting as a *tool*.

PART IV

10.0 Design Toolkit

Over the course of the year, the methods and tools utilized in this work were adapted, iterated on, and refined. In order to enable other technologists and designers to explore this domain with different communities, these methods and tools were combined into an open source design kit. Following this, an interactive survey was conducted using questions defined in the kit to assess its approachability and scalability. This survey accumulated ~600 responses from a range of demographic profiles. The kit materials, survey results, tutorial videos, and related handbook are currently being shared with the broader community through workshops and will be open sourced. This chapter describes the evolution of and final iteration of the kit.

10.1 Kit Contents

The design kit contains four key design tools: (1) action cards, (2) theme cards, (3) personality cards, and (4) design cards (Figure 10.1). Each of these components is included in its own bag with corresponding instructions. These kit box itself has a built in voice-based agent using the Google AI voice-kit and Raspberry Pi to allow all necessary components to be self contained. All tools are described in detail below.

AGENT ACTION CARDS

Agent action cards from study 1 and study 2 were iterated upon based on participant feedback and suggestions for new actions. On the basis of this feedback, agent action categories were changed to integrate ideas around technology consolidation and politeness. In addition, each action card was written as a statement with corresponding examples to allow the cards to be a

Figure 10.1: Photographs of Talking Machines Design Kit.

standalone tool without significant facilitation. The final 7 card categories included the following: (1) *functions* (i.e. alarm clock, speaker, etc.), (2) *reminders* (i.e. meetings, lists), (3) *information* (i.e. weather, news), (4) *suggestions* (i.e. workout, eating healthy), (5) *agent sharing something* (i.e. playlist, jokes), (6) *connecting with others* (i.e. phone calls, social sharing), (7) *politeness* (i.e. greetings, thank you). Blank cards were also included to allow future users to add additional actions. These card categories and example cards are shown in Figure 10.2 below.

Figure 10.2: Seven agent action card categories (colored on left) and example cards (right).

THEME CARDS

To prompt discussion surrounding voice-based agents, theme cards were developed on the basis of themes identified through study 1 and study 2. Each theme card includes a single theme and question-based prompts for the user to elicit thoughts, reflections, and ideas. 33 theme cards were designed in 5 theme categories including (1) functions an agent can help you with, (2) impact on your life, (3) role of the agent, (4) technology capabilities and use, and (5) security and privacy. Example theme cards are shown below in Figure 10.3.

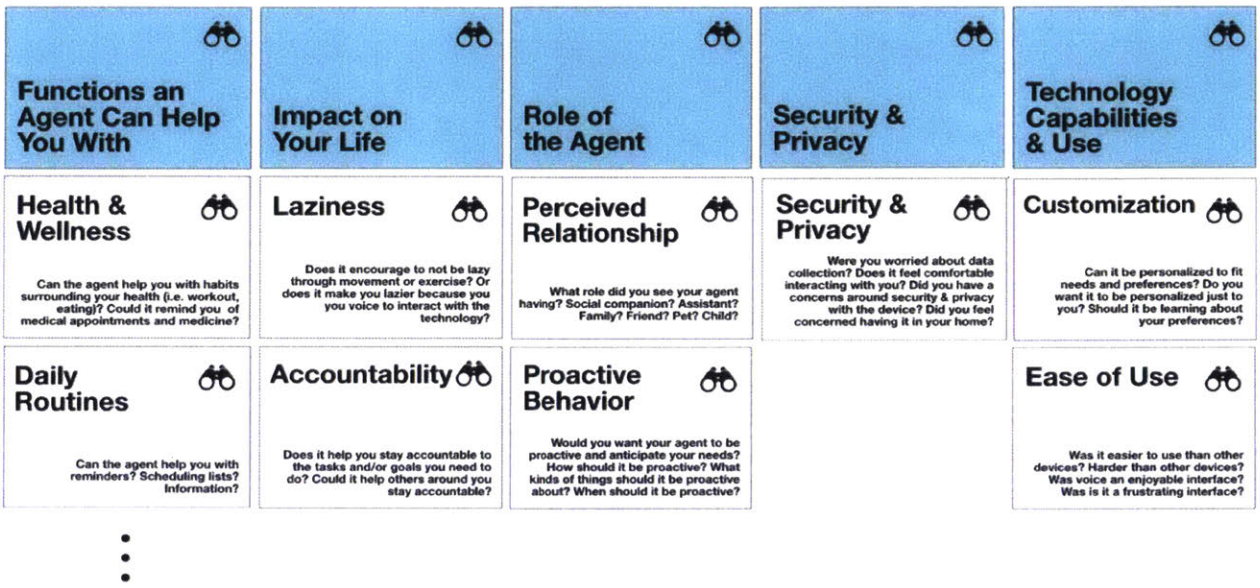


Figure 10.3: 5 theme card categories and example cards within each.

PERSONALITY CARDS

To dig into agent personality, a set of 3 personality cards were created and compiled into a small book: (1) the individual's personality, (2) their perception of an agent's personality, and (3) their dream agent's personality. Each card contains 10 personality questions and a 1-7 Likert scale. These cards are shown below in Figure 10.4.

Figure 10.4: Set of 3 personality cards for individual, agent, and dream agent.

DESIGN CARDS

The agent design cards were refined and placed in a 10-step design process as shown below including (1) form, (2) materiality, (3) inputs, (4) outputs, (5) location, (6) internet of things, (7) personality and gender, (8) ethics, and (9) rating. These cards are shown in Figure 10.5 below.

Figure 10.4: Agent design cards including 10-step design process.

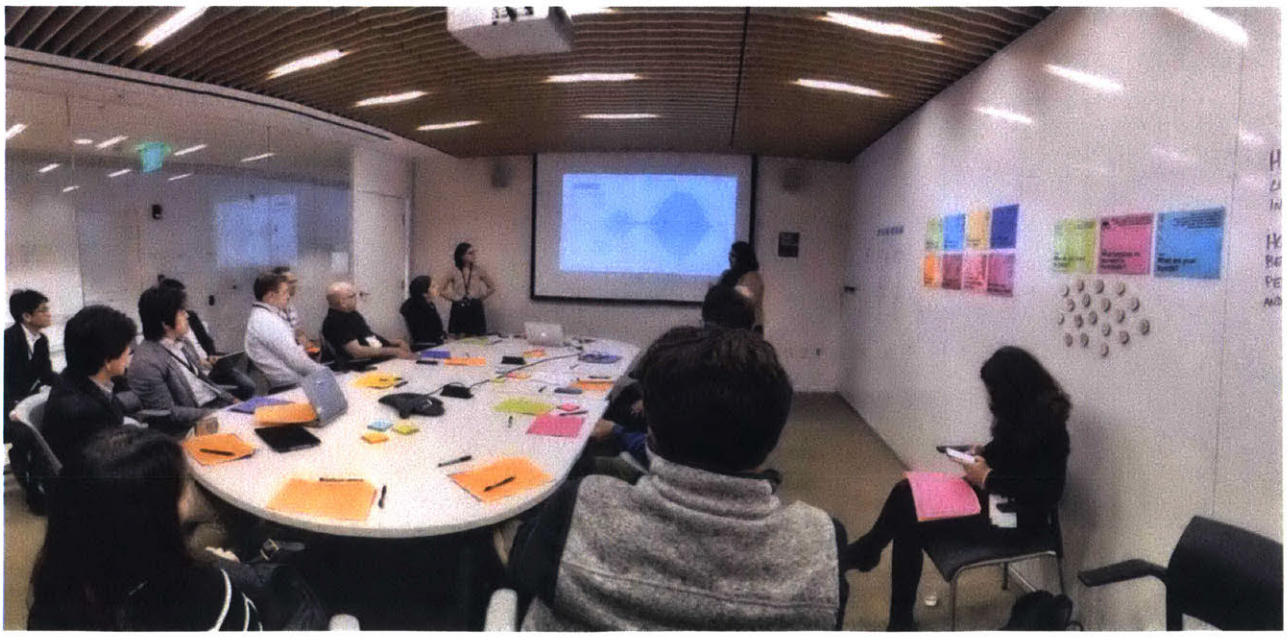
10.2 Kit Goals & Deployment

The primary goal of the kit is to provide a set of design tools to help democratize the design of voice-based agents for the home. Important aspects of the kit include:

- *Accessibility:* The kit has been tested and iterated upon with people ages 5-98. As such, it provides a "language of engagement" that resonates with a wide population.
- *Modularity:* The use of cards is modular and can be adapted to a range of problems. For example, small magnetic tokens were used in this work to explore ideas around proactivity. As such, it provides a unique means for others to explore other sub-problems within the domain.
- *Scalability:* The simplicity of the tools in the kit allows it to be printed and deployed anywhere at varying levels of complexity. The cards include simple prompts and examples that have emerged from the year-long exploration with participants. As such, these tools can be utilized by the end-user or as part of a facilitated design session.

A series of workshops using the kit have been conducted with member companies at the MIT Media Lab to further refine the tools. The kit will be shared online along with tutorial videos, a handbook, and results from a ~600-person interactive survey to enable broader access.

Figure 10.4: Agent design cards including 10-step design process.



PART IV

11.0 Contributions, Key Insights & Future Work

This thesis presented insights, learnings, and reflections from 2 studies conducted with children, adults, and older adults to understand how people interact with, discover, experience, reflect upon, and design voice-based agents for the home. The central goals of this work were to explore the following questions surrounding people's preferences, desires and boundaries for voice-based agents in the domestic environment:

7. *Interaction and Discovery:* What are people's preferences, desires and boundaries for voice-based agent technology in the home?
8. *Experience and Reflection:* How do these preferences, desires, and boundaries evolve as people live with voice-based agents?
9. *Design:* How do people envision the design of future voice-based agents for their home?

This chapter summarizes contributions, findings from this thesis as pertinent to these three core objectives and highlights open questions and areas for future work.

11.1 Contributions

This thesis explored the intersections of technology and design, people and agents, ambiguity and artificial intelligence. In doing so, the insights and learnings from this work contribute to the domains of *human-computer interaction* and *human-centered design.*

Given the emerging nature of voice-based agents and the complexities of studying the home environment, there is limited work in human-computer interaction that characterizes modern

engagement with such devices. The work represented in this thesis is thus novel in several ways.

1. **It provides an *intergenerational perspective* from children, adults, and older adults on functionalities, capabilities, and future designs of voice-based agents for the home.** While prior studies have explored voice-based technologies with targeted age groups, there is limited work that engages an intergenerational population. As such, this work is novel in its integration of a diverse population of children, adults, and older adults to explore both current and future applications of voice-based agents for the home.
2. **It reflects a *long-term engagement* of participants, allowing for rich insights surrounding in-home experiences and changes in perspective over time.** Most prior studies surrounding voice-based interfaces for the home have been conducted in the lab setting or involve short-term engagements in the home. In this work, participants were engaged over a longer period of time (6 months-1 year) to enable relationship-building and trust with facilitators and capture shifts in preferences, desires, and boundaries. In addition, participants were able to live with the agents for 1 month to allow for natural interactions and perspective building.
3. **It seeks to better characterize experiences around *living with personified artificial intelligence* in the home through the comparison of two different agent platforms (smart speaker versus social robot).** No prior work has evaluated the *affect of relational qualities, proactivity, and social expressivity* on agent perception and impact within the domestic environment. The second study in this work aims to understand and characterize these differences through a mixed-methods approach and exploration of different agent forms.
4. **It shares a series of *tools and methods* for designers and technologists to further explore and develop the space.** In addition to these contributions to the domain of human-computer interaction, this thesis also shares contributions with human-centered design. The creation and evolution of the design tools used to engage participants with voice-based agents provides insights

into building a *language of engagement* for such novel technologies. Further, the sharing of these methods and tools with the broader community through the design kit aims to help democratize the design of future voice-based agents.

11.2 Key Insights

INTERACT & DISCOVER:
WHAT ARE PEOPLE'S PREFERENCES, DESIRES, AND BOUNDARIES FOR VOICE-BASED AGENT TECHNOLOGY IN THE HOME?

Despite the proliferation of commercial voice-based agents for the home such as Amazon Echo, Google Home, and Jibo, the majority of participants had never interacted with these devices. This was particularly true in the case of older adults, who were initially often intrigued, but also skeptical of the technology. Participants with no prior exposure to using voice as an interface initially struggled with aligning their expectations with the capabilities of current voice-based systems, asking questions such as *"Alexa, what should we do about North Korea?"* or *"If there are so many of these devices around, then why have I never heard of it?"*.

Both children and adults were largely comfortable with the query-answer interaction of current voice-based agents. While children were expansive in their questions for the agents, adults tended to be more conservative in their queries. Older adults exhibited the greatest variability in their initial interactions with agents. Some older adults found interactions with voice-based agents to be natural and fun, often sharing stories or thoughts. Other older adults felt a sense of guilt in not knowing "what to ask" the agent or were confused about the purpose of the technology.

These initial impressions highlighted the need for a *language of engagement* to introduce participants to the technology and minimize barriers to entry. The creation of a card sorting activity for potential *agent actions* enabled participants to better engage with and communicate their preferences for functionalities in voice-based agents. This insight highlights the power of simple tools and methods to enable designing *with* users, instead of *for* them.

Analysis of the agent actions preferences showed that children, adults, and older adults had unique perspectives on the role they wanted voice-based agents to occupy in their lives.

Adults were the most limited in their desires, focusing on functional, more practical agent actions such as reminders, information sharing, and technology consolidation for phone calls and messages. Children, while sharing this affinity for functional aspects of the technology, were also open to the agent exhibiting social qualities. Finally, older adults exhibited the broadest openness to the technology, seeing it additionally as a tool to connect with friends and family through social sharing. Rather interestingly, all generations exhibited skepticism towards the agent providing *suggestions for wellbeing*, finding it invasive towards their independence and sense of autonomy. Qualitative analysis of participant dialogue revealed that themes surrounding the social environment, independence, agent role, agent autonomy, and privacy were central to uncovering drivers for and against socially-driven actions in the voice-based agents. Participants had the most difficulty synthesizing and sharing perspectives on their perceived relationship with the technology, emphasizing the need for designers and facilitators to design means to uncover these ideas.

EXPERIENCE & REFLECTION:
HOW DO THESE PREFERENCES, DESIRES, AND BOUNDARIES EVOLVE AS PEOPLE LIVE WITH VOICE-BASED AGENTS?

Insights from the first study highlighted the complexities of disrupting the home with technology. Participants spoke about the routines, systems and rituals that composed their home environment, often perceiving the technology as invading this personal and private ecosystem. To combat this challenge, a series of approachable and simple *probes* were designed to enable a long-term encounter with voice-based agents in the domestic environment.

Across generations, people broadly exhibited 3 unique patterns of exploration with the technology: *non-users, sporadic users, and consistent users*. Adults, echoing preferences shared in study 1, gravitated towards functional aspects of voice-based agents in their daily use. If the functional capabilities of the agent were compelling—as in the case of Amazon Alexa—adults slowly integrated social and entertainment actions into their daily usage. Children, on the other hand, were most enamored by entertainment aspects of voice-based agents. When the experience was more interactive—as in the case of Jibo—children also exhibited higher usage of functional aspects of the

technology. Finally, older adults tended to evaluate the agent as a whole rather than disambiguating significantly across action types. While older adults' usage of functional actions was consistent across agents, social and emotional interactions were more common with Jibo.

When reflecting on their experience, participants often described engaging regularly with the technology during their morning routines. Additionally, several participants discussed the habitual integration of proactively-surfaced activities like Jibo's "word-of-the-day" into their daily life. These insights, combined with data presented on usage patterns, point to ways in which designers and technologists can create *time- and context-aware, proactive interaction experiences* to re-engage users with the technology in creative ways.

After experiencing the technology for a period of 1-month, all generations exhibited shifts in their preferences, desires, and boundaries for voice-based agents in their home. While children, adults, and older adults had highly varied desires for the technology before this in-home experience, these generations became more similar in their perspectives after experiencing the technology. Participants became *more open to socially-driven aspects of the technology*, while becoming even *more conservative towards suggestions for wellbeing*.

Differences in social-expressivity and proactivity between Alexa and Jibo also led to varied perceptions of personality and role. Participants tended to perceive a more socially-expressive agent such as Jibo as more extraverted, engaging, and open to new experiences, but also less consistent and predictable. On the other hand, Alexa was seen as more consistent, dependable, and stable, but also less personable. The expressive nature of Jibo also led participants to project more feelings towards the agent. These differences were also evident in perceptions of role, wherein participants often saw Jibo as a companion, pet, child, friend or family member and Alexa more as an assistant or companion.

Though participants became more open to the role of voice-based agents in their lives after experiencing the technology, they were also highly conflicted. These conflicts were often related to projecting human qualities onto the technology, perspectives towards proactive interactions, the intended role of the agent, and concerns around security and privacy. The emergence of these concerns highlight the significant impact such technologies

can have on the home ecosystem. As such, it is critical for designers and technologists to consider these effects when building agents for the home.

DESIGN:
HOW DO PEOPLE ENVISION THE DESIGN OF FUTURE VOICE-BASED AGENTS FOR THE HOME?

Caricature and functional agent form-factors were the most commonly desired amongst participants. Participants also highlighted the need for a virtual or wearable manifestation of the agent for portability and use "on-the-go". Participants who wanted their agent to *sense* more (inputs), were also open to more modalities for *engagement* (outputs). Children desired the most multimodality in inputs and outputs while adults and older adults tended to be more conservative. Sound, vision, and motion inputs were largely accepted by most participants. On the other hand, perspectives towards the agent sensing emotion and touch were more controversial. With respect to outputs, participants often desired sound, speech and visual outputs, while movement, facial expressions, and haptic feedback tended to be less preferred. All generations tended to gravitate towards obedient and assisting and/or social and fun personalities with some differences between age groups.

Ethical stances towards voice-based agents in the home changed significantly after the in-home experience. Somewhat paradoxically, participants exhibited increased desires for *proactive interaction* in their "dream agent", but lowered willingness for data collection. Rather interestingly, participants expressed a strong desire for their dream agent to reflect their own personal world views and values. This, combined with strong desires for the agent anticipating needs and being proactive, shows that participants see voice-based agents as a technology that *caters* to the individual as opposed to acting as a *tool*.

11.3 Future Work

It is important to note, that first and foremost, the work presented in this thesis is an *exploration* of a highly complex and evolving domain. As such, while it provides key insights into people's preferences, desires, and boundaries for voice-based agents within their home, it also opens many questions that require further exploration.

Though an effort was made in this work to engage a diverse population across different age ranges, it is important to assert that age is a *single dimension* along which to evaluate preferences, desires, and boundaries for the technology. In fact, this work has shown significant variability in preferences *within* age groups as well as similarities in perspectives *across* very different age groups. This is particularly relevant in the case of older adults who represent a vast range of living styles, cognitive and physical abilities, and perspectives. As such, in future work, it is important to consider *other dimensions* surrounding individuals to holistically capture and represent perspectives towards the technology.

In this vein, another limitation of the work lies in limited geographical representation. While the crowd-sourced survey seeks to mitigate some of these effects, it is important for future work in this space to capture preferences, desires, and boundaries of more diversely located populations. In order to enable the methods and tools described in this work to be scaled effectively, partnerships with community partners are currently being explored. In addition, the design kit and related survey results will be open-sourced.

In-home deployments in this work prioritized participant comfort, privacy, and convenience to enable long-term engagement with facilitators. As such, no data was collected on devices and daily logging was only required for a period of 2 weeks. Given that prior work has shown significant changes in technology adoption after 3-6 months, further explorations should focus on embedding technologies for longer periods of time within the home environment. To enable this effectively, insights shared in this work can be leveraged to design new tools and methods that help maintain participant engagement and participation.

To enable effective contrast in interaction experiences and maintain reasonable sample sizes, only two agents (Jibo and Amazon Echo Dot) were studied in this work. Given the rapid rate of evolution of these devices and other available interaction experiences (e.g. Google Home), future work could explore a more varied set of voice-based agents. Particular focus should be placed on proactivity and social-expressiveness when conducting studies. In addition, insights derived from participants' "dream agents" designs can be utilized to prototype new forms and

interaction experiences for evaluation.

Finally, to realize the goal of democratizing the design of voice-based agents for the home, it is critical to make the learnings, insights, and reflections from this work accessible to a wider population. The objectives in doing this should not be solely to share data and tools, but also to provoke people to consider the impact of these technologies on their lives.

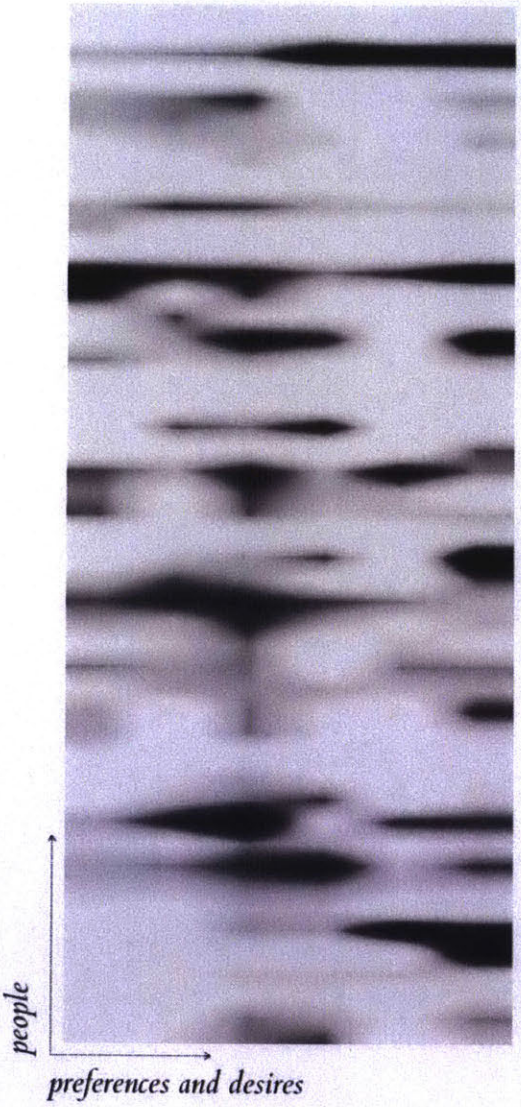


Figure 11.1: This heat map of raw data on agent action preferences from all participants in this work aims to show the many *shades of gray* that exist in people's preferences, desires, boundaries, and perceptions of the technology. It is thus our responsibility to design *for* and *with* people as we create the *human-agent relationship.*

To the 69 people who took part in this journey, thank you for *talking machines* with us.

12.0 References

1. Bickmore, Timothy W, and Rosalind W Picard. 2005. "Establishing and Maintaining Long-Term Human-Computer Relationships." *ACM Transactions on Computer-Human Interaction* 12(2): 293–327.

2. Bickmore, Timothy, Daniel Schulman, and Langxuan Yin. 2010. "Maintaining Engagement in Long-Term Interventions with Relational Agents." *Applied Artificial Intelligence* 24(6): 648–66.

3. Branco, R. M., Quental, J., & Ribeiro, Ó. (2017). Personalised participation: an approach to involve people with dementia and their families in a participatory design project. *CoDesign*, *13*(2), 127-143.

4. Brandt, Eva, and Jörn Messeter. 2004. "Facilitating Collaboration through Design Games." *Proceedings of the Eighth Conference on Participatory Design Artful Integration: Interweaving Media, Materials and Practices - PDC 04* 1: 121.

5. Breazeal, Cynthia. 2003. "Toward Sociable Robots." *Robotics and Autonomous Systems* 42(3–4): 167–75.

6. Brennan, S. E. (1998). The grounding problem in conversations with and through computers. *Social and cognitive approaches to interpersonal communication*, 201-225

7. Brennan, Susan E. 1991. "Conversation with and through Computers." *User Modeling and User-adapted Interaction* 1(1): 67–86.

8. Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., & Yan, H. (1999, May). Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 520-527). ACM.

9. Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., & Yan, H. (2001). More than just a pretty face: conversational protocols and the affordances of embodiment. *Knowledge-based systems*, *14*(1-2), 55-64.

10. Cassell, Justine. 2000. "Embodied Conversational Interface Agents." *Communications of the ACM* 43(4): 70–78.

11. Cha, Elizabeth, Jodi Forlizzi, and Siddhartha S Srinivasa. 2015. "Robots in the Home: Qualitative and Quantitative Insights into Kitchen Organization." *Hri*: 319–26.

12. R., Yilmaz, S., Seifert, C., & Gonzalez, R. (2012). Design heuristics support two modes of idea generation: Initiating ideas and transitioning among concepts

13. Crabtree, Andy, and Tom Rodden. 2004. "Domestic Routines and Design for the Home." *Computer Supported Cooperative Work: CSCW: An International Journal* 13(2): 191–220.

14. Cramer, Henriette, Nicander a. Kemper, Alia Amin, and Vanessa Evers. 2009. "The Effects of Robot Touch and Proactive Behavior on Perceptions of Human-Robot Interactions." *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction - HRI '09*: 275–76.

15. Creswell, J W. 2013. Research design Qualitative quantitative and mixed methods approaches *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.*

16. Daly, Shanna R., Colleen M. Seifert, Seda Yilmaz, and Richard Gonzalez. 2016. "Comparing Ideation Techniques for Beginning Designers." *Journal of Mechanical Design* 138(10): 101108.

17. Dautenhahn, Kerstin et al. 2005. "What Is a Robot Companion - Friend, Assistant or Butler?" *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*: 1488–93.

18. De Graaf, Maartje, Somaya Ben Allouch, and Jan van Dijk. 2017. "Why Do They Refuse to Use My Robot?" *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction- HRI '17*: 224–33.

19. Desjardins, A., Wakkary, R., & Odom, W. (2015, April). Investigating genres and perspectives in HCI research on the home. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3073-3082). ACM.

20. DeVault, David et al. 2014. "SimSensei Kiosk : A Virtual Human Interviewer for Healthcare Decision Support." *International Conference on Autonomous Agents and Multi-Agent Systems* (1): 1061–68.

21. Druin, Allison. 1999. "Cooperative Inquiry: Developing New Technologies for Children with Children." *Human Factors in Computing Systems (CHI)*: 592–99.

22. Editor, Isabel L Nunes. 2018. 592 *Advances in Human Factors and Systems Interaction.*

23. Ehn, Pelle. 2008. "Participation in Design Things." *Conference on Participatory Design*: 92–101

24. Frohlich, D. M. (1997). Direct manipulation and other lessons. In *Handbook of Human-Computer Interaction (Second Edition)* (pp. 463-488).

25. Gaver, W.W., J. Beaver, and S. Benford. 2003. "Ambiguity as a Resource for Design." *SIGCHI conference on Human factors in computing systems* (5): 233–240.

26. Gaver, William W. et al. 2013. "Indoor Weather Stations: Investigating a Ludic Approach to Environmental HCI Through Batch Prototyping." *CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 3451–60.

27. Gosling, Samuel D., Peter J. Rentfrow, and William B. Swann. 2003. "A Very Brief Measure of the Big-Five Personality Domains." *Journal of Research in Personality* 37(6): 504–28.

28. Halskov, Kim, and Peter Dalsgård. 2006. "Inspiration Card Workshops." *Proceedings of the 6th ACM conference on Designing Interactive systems- DIS '06*: 2.

29. Hanington, B, and B Martin. 2012. Develop Innovative Ideas *Universal Methods of Design: 100 Ways to Research Complex Problems.*

30. Hutchinson, Hilary et al. 2003. "Technology Probes: Inspiring Design for and with Families." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)* (5): 17–24.

31. Kachouie, Reza, Sima Sedighadeli, Rajiv Khosla, and Mei Tai Chu. 2014. "Socially Assistive Robots in Elderly Care: A Mixed-Method Systematic Literature Review." *International Journal of Human-Computer Interaction* 30(5): 369–93.

32. Kensing, Finn, and Jeanette Blomberg. 1998. "Participatory Design: Issues and Concerns." *Computer Supported Cooperative Work (CSCW)* 7(3–4): 167–85.

33. Kensing, Keld Bødker Finn, and Jesper Simonsen. 2004. "Participatory IT Design."

34. Klamer, T., Allouch, S. B., & Heylen, D. (2010, June). "Adventures of Harvey"–Use, acceptance of and relationship building with a social robot in a domestic environment. In *International Conference on Human-Robot Personal Relationship* (pp. 74-82). Springer, Berlin, Heidelberg

35. Kopp, Stefan, Lars Gesellensetter, Nicole C. Krämer, and Ipke Wachsmuth. 2005. "A Conversational Agent as Museum Guide - Design and Evaluation of a Real-World Application." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 3661 LNAI: 329–43.

36. Landis, J Richard, and Gary G Koch. 2016. "The Measurement of Observer Agreement for Categorical Data Published by : International Biometric Society Stable" 33(1): 159–74.

37. Leppänen, Sanna, and Marika Jokinen. 2003. "Daily Routines and Means of Communication in a Smart Home." In *Inside the Smart Home*, 207–25.

38. Licklider, J. C. R. 1960. "Man-Computer Symbiosis." *IRE Transactions on Human Factors in Electronics* HFE-1(1): 4–11.

39. Luger, Ewa, and Abigail Sellen. 2016. "'Like Having a Really Bad PA': The Gulf between User Expectation and Experience of Conversational Agents." *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*: 5286–97.

40. Luria, Michal et al. 2016. "Designing Vyo, a Robotic Smart Home Assistant: Bridging the Gap between Device and Social Agent." *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*: 1019–25.

41. Luria, Michal, Guy Hoffman, and Oren Zuckerman. 2017. "Comparing Social Robot, Screen and Voice Interfaces for Smart-Home Control."

42. Martin, C.D. 1997. "The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places [Book Review]." *IEEE Spectrum* 34(3): 9–10.

43. Massung, Elaine et al. 2015. "Ambiguity as a Resource for Design." *SIGCHI conference on Human factors in computing systems* 13(5): 233–240.

44. McTear, Michael, Zoraida Callejas, and David Griol. 2016. The Conversational Interface *Creating a Conversational Interface Using Chatbot Technology.*

45. McTear, Michael, Zoraida Callejas, and David Griol. 2016. The Conversational Interface *Evaluating the Conversational Interface.*

46. Monk, Andrew. "Notes Towards an Ethnography": 276–81.

47. Moore, R. K. (2014, October). Spoken Language Processing: Time to Look Outside?. In *International Conference on Statistical Language and Speech Processing* (pp. 21-36). Springer, Cham.

48. Muller, M.J. 2003. "Participatory Design: The Third Space in HCI." *Human-Computer Interaction Handbook* 4235: 1051–68.

49. Kuhn, Sarah, and Michael J. Muller. "Participatory design." *Communications of the ACM* 36.6 (1993): 24-29.

50. Ogonowski, Corinna et al. 2013. "Designing for the Living Room: Long-Term Users Involvement in a Living Lab." *Chi 2013*: 1539.

51. Patton, Michael Quinn. 2014. Thousand Oaks, California *Qualitative Research & Evaluation Methods. 2001 Sage Publications.*

52. Peltu, Malcolm. 2008. "Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues." *Forum American Bar Association* (14): 1–33.

53. Porcheron, Martin, Joel E Fischer, Stuart Reeves, and Sarah Sharples. 2018. "Voice Interfaces in Everyday Life." *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems.*

54. Portet, François et al. 2013. "Design and Evaluation of a Smart Home Voice Interface for the Elderly: Acceptability and Objection Aspects." *Personal and Ubiquitous Computing* 17(1): 127–44.

55. Gaver, B., Dunne, T., & Pacenti, E. (1999). Design: cultural probes. *interactions*, *6*(1), 21-29.

56. Qvarfordt, P. 2004. "Eyes on Multimodal Interaction.". Univ. 2004.

57. Rogers, Y. (2006, September). Moving on from Weiser's Vision of Calm Computing: Engaging Ubicomp Experiences. In *International conference on Ubiquitous computing* (pp. 404-421). Springer, Berlin, Heidelberg.

58. Sanders, E. B. N., & Stappers, P. J. (2014). Probes, Toolkits and Prototypes: Three Approaches to Making in Codesigning. *CoDesign*, *10*(1), 5-14.

59. Sanders, E. B. N. (2003). From user-centered to participatory design approaches. In *Design and the social sciences* (pp. 18-25). CRC Press.

60. Sathi, A. (2016). *Cognitive (Internet of) Things: Collaboration to Optimize Action.* Springer.

61. Seale, Jane, Claudine Mccreadie, Alan Turner-Smith, and Anthea Tinker. 2002. "Older People as Partners in Assistive Technology Research: The Use of Focus Groups in the Design Process." *Technology and Disability* 14: 21–29.

62. Shechtman, Nicole, and L.M. Horowitz. 2003. "Media Inequality in Conversation: How People Behave Differently When Interacting with Computers and People." *In Proceedings of the SIGCHI conference on Human factors in computing systems*: 281–288.

63. Shneiderman, Ben. 1982. "The Future of Interactive Systems and the Emergence of Direct Manipulation." *Behaviour and Information Technology* 1(3): 237–56.

64. Sung, J, H I Christensen, and R E Grinter. 2009. "Robots in the Wild: Understanding Long-Term Use." *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*: 45–52.

65. Threatt, Anthony L. et al. 2014. "An Assistive Robotic Table for Older and Post-Stroke Adults." *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*: 673–82.

66. Torta, Elena et al. 2012. "Attitudes Towards Socially Assistive Robots in Intelligent Homes: Results From Laboratory Studies and Field Trials." *Journal of Human-Robot Interaction* 1(2): 76–99.

67. Torta, Elena et al. 2014. "Evaluation of a Small Socially-Assistive Humanoid Robot in Intelligent Homes for the Care of the Elderly." *Journal of Intelligent and Robotic Systems: Theory and Applications* 76(1): 57–71.

68. Trappl, Robert, and Andrew Patrick. 2013. Human Factors of Trustworthy Agents *Your Virtual Butler: The Making-Of*.

69. Vandemeulebroucke, Tijs, Bernadette Dierckx de Casterlé, and Chris Gastmans. 2018. "The Use of Care Robots in Aged Care: A Systematic Review of Argument-Based Ethics Literature." *Archives of Gerontology and Geriatrics* 74(September 2017): 15–25.

70. Von Der Pütten, Astrid M., Nicole C. Krämer, Jonathan Gratch, and Sin Hwa Kang. 2010. "'It Doesn't Matter What You are!' Explaining Social Effects of Agents and Avatars." *Computers in Human Behavior* 26(6): 1641–50.

71. Weiser, M. 1991. "The Computer for the 21st Century." *Scientific American (International Edition)* 265(3): 66–75.

72. Wilson, Charlie, Tom Hargreaves, and Richard Hauxwell-Baldwin. 2015. "Smart Homes and Their Users: A Systematic Analysis and Key Challenges." *Personal and Ubiquitous Computing* 19(2): 463–76.

73. Woodruff, Allison, Sally Augustin, and Brooke Foucault. 2007. "Sabbath Day Home Automation: 'It's Like Mixing Technology and Religion.'" *CHI '07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 527–36.

74. Yilmaz, Seda, Colleen Seifert, Shanna R. Daly, and Richard Gonzalez. 2016. "Design Heuristics in Innovative Products." *Journal of Mechanical Design* 138(7): 71102.

75. Yip, Jason et al. 2013. "Brownies or Bags-of-Stuff? Domain Expertise in Cooperative Inquiry with Children." *Proceedings of the 12th International Conference on Interaction Design and Children - IDC '13*: 201–10.