"Siri Talks at You": An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind

Ali Abdolrahmani
UMBC
Baltimore, MD
aliab1@umbc.edu

Ravi Kuber
UMBC
Baltimore, MD
rkuber@umbc.edu

Stacy M. Branham
UC Irvine
Irvine, CA
sbranham@uci.edu

ABSTRACT
Voice-activated personal assistants (VAPAs)—like Amazon Echo or Apple Siri—offer considerable promise to individuals who are blind due to widespread adoption of these non-visual interaction platforms. However, studies have yet to focus on the ways in which these technologies are used by individuals who are blind, along with whether barriers are encountered during the process of interaction. To address this gap, we interviewed fourteen legally-blind adults with experience of home and/or mobile-based VAPAs. While participants appreciated the access VAPAs provided to inaccessible applications and services, they faced challenges relating to the input, responses from VAPAs, and control of information presented. User behavior varied depending on the situation or context of the interaction. Implications for design are suggested to support inclusivity when interacting with VAPAs. These include accounting for privacy and situational factors in design, examining ways to support concerns over trust, and synchronizing presentation of visual and non-visual cues.

Author Keywords
Voice Activated Personal Assistant; VAPA; Blind Individuals; Usability Challenges; Non-Visual Interaction.

ACM Classification Keywords
K.4.2. [Computers and society]: Social issues – assistive technologies for persons with disabilities.

INTRODUCTION
Recent advances in technology have led to the development of voice-activated personal assistants (VAPAs) for both home environments (e.g., Google Home, Amazon Echo), and for use while on-the-go (e.g., Apple Siri, Microsoft Cortana, in-car assistants). Tasks such as creating alerts/setting reminders, searching for information, requesting music to be played or directions to be presented, and purchasing items can be conducted through voice input. A key characteristic of VAPAs (also termed as Conversational Agents, Virtual Personal Assistants, Voice-Enabled Assistants or Intelligent Personal Assistants [8]) is their capability to react and respond to the user's requests, to engage in human-like behavior and, even to be able to learn and evolve [17]. Voice-based interaction offers a practical solution to perform secondary tasks when the user is engaged with an attention-demanding primary task (e.g., looking up next steps of a recipe while kneading bread), and provides a more usable alternative to other traditional methods of input (e.g., through a virtual keyboard [2]).

While researchers have examined and analyzed the types of requests made to VAPAs [8,10,11,16] and examined issues relating to the constraints and acceptability of these technologies [9,15], studies have yet to examine VAPA usage by blind individuals and the barriers which they may face. A recent study by Pradhan et al. [21] analyzing the reviews of verified purchases of Amazon Echo VAPAs revealed that nearly 38% of reviews mentioning disabilities related to individuals with visual impairments or blindness, suggesting that these types of non-visual interaction platforms may prove useful to this community. Voice-based interaction can provide a more efficient alternative to entering text through a (virtual) keyboard and checking input using an assistive technology (e.g., screen reader) [2], One unique advantage of using VAPAs is that individuals who are blind can interact with a range of technologies which would otherwise be difficult to control due to their lack of accessibility (e.g., digital touch screen thermostats).

In this paper, we describe an exploratory study designed to understand how individuals who are blind experience VAPAs and the challenges which are encountered when attempting to access these technologies. A set of semi-structured interviews were conducted with fourteen legally-blind VAPA users who normally rely on screen reading solutions to access computing and mobile interface content. The findings demonstrate the frustrations associated with inputting commands, the presentation of responses from VAPAs, and control of this information. The impact of situational factors was also found to play a role in the ways users may interact with VAPAs due to issues of privacy. This research represents the first step in a larger project,

ASSETS '18, October 22–24, 2018, Galway, Ireland

ACM ISBN 978-1-4503-5650-3/18/10…$15.00
https://doi.org/10.1145/3234695.3236344

with the long-term goal of developing guidance to better support interface developers when attempting to design VAPA-based technologies for a more diverse audience.

RELATED WORK

Interaction with VAPAs

With a range of smart speakers, mobile voice user interfaces, and third party smartphone tools (e.g., Vlingo, Maluuba, and Evi [9]) available for use, VAPAs are becoming more commonly utilized for performing tasks while at home or while on-the-go. Though there is a growing body of research on these types of interfaces, there is still limited awareness of how such conversational interfaces are used in everyday life, especially for people who are blind.

Research suggests that VAPAs are most commonly used for queries such as determining environmental conditions [16], to listen to the news, for playing music, and controlling external devices [15]. If queries or commands cannot be understood by VAPAs, users have been found to progressively simplify these until the system is able to provide an appropriate response [16]. While accuracy of speech recognition by tools such as Siri has improved over time, it still takes significant cognitive effort to think about how to phrase commands/queries [8]. Research has examined VAPA technology adoption, including characteristics that change users' adoption resistance, the preferred user interfaces to interact with different services provided in different contexts, and user perception and experience with technology [6,7,12]. Chen et al. [6] reported that "low transparency and poor explainability" of the mechanism inside VAPAs impacts on individual's understanding of VAPAs' capabilities, and as a result impacts the way users interact with VAPAs. Findings from the study by Lopatovska et al. [15] suggest that people are satisfied with home-based VAPAs even when they do not produce sought information. This suggests that the interaction experience is more important to the users than the interaction output.

Interaction style varies depending on VAPA platform and context of use. One study found that Google Home elicited more "serious" interactions intended to accomplish work, while Apple Siri elicited more conversational, playful interactions intended to trigger inbuilt humorous responses [16]. Another study found usage often varied by location and the presence of third parties [8]. Fears of social embarrassment were highlighted when using these technologies in public spaces [8]. Participants were also found to be cautious of disclosing private information, particularly when using VAPAs in public environments (e.g., [9,10]).

Voice Interaction for Diverse Groups

Researchers have begun to examine the perceptions and needs of groups with differing needs to mainstream users. Examples include the study by Wulf et al. [25], which observed older adults using Siri. While participants were noted to appreciate the simplicity of voice-based interaction, some challenges were identified in practical usage (e.g., with initiating the speech interaction or unintentionally interrupting the interaction). Lack of consistency in the output from Siri was also found to negatively impact user behavior. In a separate study, Portet et al. [20] recruited eight older adults along with their relatives and caregivers, to participate in both interviews and observations interacting with a simulated prototype for a smart home. The voice interface was thought to offer potential to support older and frail adults, and was found to be better accepted compared to more intrusive solutions. However, only half of the older adult group found it natural to interact by voice, even though some admitted speaking aloud when they are alone. In their comparative study of older, middle-age, and young age groups, Schiavo et al. [22] explored how members of these groups use mobile multimodal interaction in everyday activities (i.e. taking photos with a tablet) by using mid-air gestures and voice commands. Investigating commonalities and differences between the groups revealed that while all groups understood how to use both modalities in combination, the older groups used gesture and voice commands in a similar way. While valuable, these studies did not specifically focus on the visual impairments faced by some older adults, which may impact interaction with these types of technology.

The challenges faced by individuals who are blind when interacting with speech-output technologies have been extensively documented by researchers (e.g., when using screen readers [14,19]). While less is known about the experiences and barriers when accessing VAPAs, studies have examined issues associated with speech input for individuals who are blind, in terms of mobile device usage [2,26] for purposes of navigation [13] and supporting education [3,18,23,24]. In terms of mobile devices, Azenkot et al. [2] conducted a survey with blind and sighted participants to investigate how often, what for, and why individuals who are blind use speech for input. Levels of satisfaction with speech were higher among the blind group, which was most likely attributed to keyboard input with VoiceOver being slower than standard keyboard input. A subsequent study examined speech input on an iPod compared with the on-screen keyboard with VoiceOver. Findings showed that even when using speech input, 80% of participants' time was spent editing the text output by the recognizer. In order to support accessible interactions with a mobile device, Zhong et al. [26] have described the design of JustInSpeak. This solution offers two main benefits for individuals who are blind. Firstly, the technology, designed for the Android OS, can reduce the time and effort needed to access application controls, as the user needs to simply voice the application that he/she wishes to open. Secondly, the solution does not interfere with existing screen readers, but works with them instead [26].

The study described in this paper focuses on the ways in which VAPAs are used by individuals who are blind, providing insight into perceptions and attitudes towards these technologies. In contrast to related work, our study represents one of the first initiatives examining experiences with both home and mobile VAPAs. While sighted users are known to view these interfaces as "entertaining / gimmicky" [16,21], these technologies are considered serious tools for blind users [2]. Blind and sighted individuals are known to differ in terms of preferences regarding speed and non-visual nature of voice user interface technologies [5]. They experience differing challenges related to privacy and safety [4]. The study described in this paper extends this growing area of accessibility research by providing deeper insight into the contexts where VAPAs are used, and the challenges faced, in order to identify more inclusive VAPA design opportunities.

STUDY DESIGN

The goal of our exploratory study was to elicit a diversity of viewpoints and practices surrounding VAPAs. Thus, a semi-structured interview protocol was used for data collection. Participants were encouraged to share their experiences with interaction, identify challenges faced, and describe the impact of settings (e.g., public/home settings, presence of others) on their experience.

Study Participants

Fourteen participants (9 male, 5 female, aged 21 to 66, mean: 31) were recruited through a combination of mailing lists attracting individuals who are blind, and snowball sampling (Table 1). Our research question required recruitment of blind users of both mobile and home VAPAs. Being relatively new technologies, this attracted early adopters; a phenomenon not uncommon among accessibility research studies.

All participants identified themselves as legally-blind with little to no residual vision. All utilized screen readers as their primary tool for accessing computing and mobile technologies (13 iPhone users, 1 Android OS user). Examples of assistive apps used on their mobile devices included TapTapSee, Nearby Explorer, and Microsoft Seeing AI. All participants used VAPAs such as Siri or Google Assistant on their mobile devices on a daily or weekly basis. Eight participants owned an Amazon Echo and/or a Google Home device, while four had experience using these technologies when visiting family or significant others. All stated they use home-based VAPAs multiple times per week, with many using them on a daily basis.

Interviews and Data Analysis

The semi-structured interview protocol consisted of three sections: interaction experiences with VAPAs, accessibility of these tools, and perceptions of privacy/security. Probing questions were used to stimulate discussion to elicit data in different situations and contexts (e.g., in public/home settings and in presence of strangers or friends and family members). Interviews were conducted remotely using audio/video communication software (e.g., Skype and Facetime Audio). These were audio-recorded. Each interview lasted between 40 and 80 minutes (mean: 55 minutes), and participants were paid $20 for their time. Interview recordings were transcribed verbatim by two researchers.

The primary author used inductive thematic analysis to identify themes from participant responses. A second researcher independently coded a random subset of the interviews. This analysis showed strong inter-coder agreement between the two researchers (Cohen's Kappa coefficient (κ) = 0.76).

FINDINGS

Scenarios of Usage

Participants described using VAPAs to perform an array of tasks, from setting reminders to providing access to apps, services, and devices (Table 2). Reasons for using VAPAs broadly related to their convenience for addressing queries and performing commands. When asked about instances when these technologies would be used, participants suggested scenarios where a timely response was needed (e.g., asking Siri "*Where am I?*" while navigating to reduce time taken asking others for assistance). If the user was engaged with another task, VAPA interaction was thought to promote the ability to multi-task (e.g., asking Siri for definition/synonyms of words or to perform a mathematical conversion while writing a report at work), or in cases where one's hands were occupied (e.g., asking Google Home to dictate a recipe while cooking), among others. Voice interaction was described as a "*time saver*," reducing task time compared to inputting data through touch screens or using a keyboard. It also supported access to technologies that would otherwise be inaccessible (e.g., controlling specific digital home thermostats), enabling participants to perform tasks independently without the need for sighted assistance or specialized software/hardware.

Usability and Accessibility of VAPAs

Misinterpreting Input Commands

When asked about their experiences with VAPAs, ten participants highlighted that mobile VAPAs such as Siri were found to misinterpret their commands or queries, particularly in public venues when ambient sounds may mask or interfere with recognition (e.g., when riding on a noisy train). Participants found it frustrating when an unexpected action was made by a VAPA due to misinterpretation of the original command. This would lead to lack of confidence in VAPAs to perform tasks such as setting reminders while in public settings (P1). P1 highlighted that she would rarely ask Siri for details of upcoming reminders, as the output was frequently incorrect. Instead, she would wait for a notification to appear on her iPhone, or check it manually on the long list of her reminders. P8 had given up using Siri to set reminders or

ID	Age	Gender	Visual Impairment Status	Mobile VAPA	Mobile VAPA Frequency of Use	Home VAPA	Home VAPA Frequency of Use
P1	27	F	From birth	Siri	SD	Echo*	FW
P2	36	F	Later in life	Siri	SD	Echo**	SD
P3	24	F	From birth	Siri	OD	N/A	N/A
P4	37	M	Later in life	Siri	SD	Echo*	SD
P5	33	M	From birth	Siri, Google Assistant	SD, SD	Google Home**	SD
P6	25	F	From birth	Siri	SD	N/A	N/A
P7	35	M	From birth	Siri	OD	Echo**	SD
P8	25	M	From birth	Siri	FW	Echo**	SD
P9	24	M	From birth	Siri, Google Assistant	SD, FW	Echo**	SD
P10	22	M	Later in life	Siri, Google Assistant	SD, OD	Echo*	SD
P11	38	M	From birth	Google Assistant	SD	Echo**, Google Home**	SD, SD
P12	66	M	From birth	Siri	FW	Echo**, Google Home**	SD, SD
P13	21	M	From birth	Siri	SD	Echo**	SD
P14	27	F	From birth	Siri	OD	Echo*	FW

Table 1. Participant demographics. VAPA frequency of use: SD: several times a day, OD: once a day, FW: a few times a week, OW: once a week, R: rarely. Home VAPA: *used, **owned.

calendar events, as he found he needed to spend too much time checking for errors. It was faster to enter important events through keystrokes in the calendar app using his touch screen.

Other challenges included difficulty recognizing accents or the colloquial phrases used when speaking (P7). P7 was surprised that Siri "*could not learn over time*." P13 described difficulties with Siri misinterpreting names from his contact list. He stated that sometimes Siri would get confused, and even repeating the correct name clearly several times did not lead to the correct contact being called. P14 could not find a way to have Siri spell out her name correctly in text conversations. She found Siri to be weak in composing long messages that include names. She would therefore use Siri only for composing short, simple text messages.

Challenges recognizing names and interpreting complex commands were not only confined to mobile VAPAs. Three participants reported similar issues when using home-based VAPAs. The difficulties were thought to be exacerbated when the user was at a farther distance from the VAPA, or when voicing a command while music was playing through the device. Lack of faith in VAPAs performing tasks appropriately led to participants developing strategies to verify that commands voiced had been addressed. For example, P7 described creating appointments using Siri rather than through the Calendar app (similar to P8). He would always double check the details prior to sending invitations to his colleagues or clients at work to avoid embarrassing mistakes: "*If I send an invitation to somebody, it'll send everything I dictated, and if that's not correct, it'll come across as unprofessional*."

Categories	Examples	Home	Mobile
Managing time	Setting timers/alarms, setting reminders, and calendar appointments	8	14
Information seeking	Determining time and weather, gathering directions, checking facts	10	14
Access to apps, services and devices	Making phone calls, text messaging, home device control, third party apps (Uber, Spotify, etc.)	5	14
Media and entertainment	Playing music, listening to the news, playing games and listening to jokes	12	7
Math and language aid	Supporting basic arithmetic and conversions, word definitions and synonyms, and sentence translations	1	8

Table 2. Frequency of participants performing tasks using home and mobile VAPAs.

In order to input a voice command, users must activate VAPAs by either physically pressing a button on the device or providing a verbal cue (e.g., preceding each command with "OK Google"). Upon activation, a specific time period is allocated to begin and complete the process of voicing a command. Four participants found this timeout frustrating, especially when inputting a complex command, such as making a calendar appointment or composing an email. Mention was made of there not being enough feedback provided to estimate the time window available, prior to the system timing-out. P6 indicated, "*You have to be pretty quick with that one [the calendar], otherwise it'll cut you off.*"

Feedback from VAPAs

Six participants shared experiences where feedback was too verbose, unnecessary, irrelevant, or conversely, provided insufficient information to satisfy requests made. In terms of verbosity, P1 described being in her workplace and asking Siri to call her boyfriend. She whispered his name into the phone, as she did not want to draw attention about making personal calls during work time. In response to her command Siri loudly voiced his name, revealing to her coworkers that she was making a personal call, and identifying the person being dialed. She highlighted that she could not customize the amount of feedback presented for certain tasks. She hypothesized that if Siri could say "*OK*" to confirm that a contact was being dialed, instead of revealing the full name of the contact, it would support performing tasks more discreetly.

Another challenge related to presenting extraneous feedback in response to voice commands. When attempting to use Siri to identify opening times of a restaurant she wanted to go to, P1 stated that the address of the restaurant was presented, instead of a concise response just describing times. As a workaround, P1 had to use the Google Maps app on her phone to find the answer. P14 serendipitously learned that after asking Siri an initial question (e.g., "*which restaurants are near me*?"), follow-up questions like "*is it open now*?" may lead to the response she was looking for. Other participants were largely unaware of this functionality. She noted that sighted users may face similar problems when using Siri, but they have alternative options for readily accessing this information; if they use traditional tools (e.g., Google Maps) to perform tasks, they may be able to quickly glance at the screen to identify areas of interest on the screen relating to their query, rather than listening to content sequentially.

Supporting Apps

VAPAs were thought to offer considerable potential to blind users in particular by enabling them to interact with third party applications and services (e.g., arranging rideshares using Siri, accessing the Spotify music service using Amazon Echo). VAPAs allowed users to perform a variety of tasks faster and reduce effort needed to interact with these services or apps, some of which are screen reader inaccessible, as compared to using their phone's touch screen. However, three participants described the need for improvement, suggesting that Siri in particular should support voice access to more apps and services (e.g., Spotify and Google Calendar). Suggestions were also made to support more complex tasks through voice interaction. For example, P6 highlighted that interaction for composing emails should be more straightforward, without the pressure of attempting to perform tasks "*before Siri times-out*." A general theme identified from participants related to the ways in which small improvements to a VAPA could substantially improve the quality of the user experience.

Recovering from Speech Recognition Errors

The process of both identifying errors and then recovering from them was found to be challenging. P6 highlighted that with VoiceOver enabled, Siri reads back the dictated message prior to sending it out to ensure that messages are free of mistakes. Her frustration arose from the time and cognitive effort needed to identify and correct errors in long messages, because to fix an error in a long message, the entire message should be dictated again. Two other participants described difficulties with Siri, particularly in instances when recovering from errors related to more complex voice commands (e.g., setting a recurring reminder or calendar appointment).

Adjusting Voice Settings on VAPAs

Advances in text-to-speech technologies have led to the design of speech feedback that sounds more natural to users, as compared with robotic-sounding feedback when speech-synthesizers were first introduced. While the improvements in quality of speech were appreciated by participants, mention was made that VAPAs should provide the ability to allow users to customize voice output settings (e.g., speech rate, clarity, and intensity) based on their needs and the nature of the task. Experienced screen reader users are accustomed to speech presented via their devices at a fast rate. However, VAPAs caused frustration when the user was forced to interact at a slower pace than desired, as the speed of information presented could not be easily controlled (P9). Additionally, five participants indicated that faster presentation of content would help them preserve their privacy when interacting with VAPAs without headsets, particularly when in public or in the presence of others. They reasoned that the fast output would sound unintelligible to most sighted passersby.

Voice Activation "Hey Siri, Alexa, Hey Google"

VAPAs can be activated through 'wakewords' such as "*Hey Siri*." However, eight participants reported instances of their VAPAs not responding when wakewords were voiced, or conversely were unexpectedly activated. Four participants indicated that Siri on iPhones was the largest offender. Their experiences had led to them favoring physically-activating Siri, rather than using a wakeword, to ensure that they could stay in control of activation. Both P10 and P14 described concerns about the unexpected activation of their VAPAs in social settings. To reduce the risk of unintentional activation, P10 would keep his phone in his pocket when in public to muffle the microphone. P14 disabled the wakeword feature altogether to prevent Siri from taking undesired actions out of her control.

Participants found that this challenge would occur more frequently with their home VAPAs, particularly if the television was switched on and names or words akin to "Alexa" were presented. Three participants found Amazon Echo's microphone to be so sensitive, that they would switch it off in certain situations such as social gatherings. For some, concern stemmed from hearing stories of random items being added to one's Amazon shopping cart, which might then be accidentally purchased.

Visual Content Creating Usability Challenges

P8 encountered difficulties identifying system status regarding whether the microphone was on or off. He stated, "*It took a while to figure out what the button [on the Echo device] did. It is tactile but does not say what it does. It does make noise when you press it. My girlfriend found it because she has a small level of sight available to her. I guess if you weren't aware of that feature, the button could be confusing.*" As a workaround, P8 would test if his Echo was on and the microphone was working by using the wakeword "Alexa." If the VAPA would not respond, he would assume that the microphone was switched off.

The same participant highlighted other issues where visual cues were not presented effectively through alternative means. Examples include the Amazon Echo light ring, which blinks to indicate a reminder notification and changes color to indicate muting of the microphone. He remarked that this could be frustrating for blind users, as they were not privy to the same cues as their sighted counterparts. Furthermore, in situations where sighted users may be in the same environment, it seemed unfair that they would be able to detect these signals, particularly when troubleshooting problems, prior to blind users. This was thought to lead to reliance on sighted users for purposes of assistance, when the technology itself should really be accessible to all. It could also contribute to blind users' privacy being compromised, as some users may feel inclined to have sighted assistance on-hand.

Comparison of VAPA Platforms

Participants who had used multiple VAPAs were able to compare and contrast their experiences, describing these in terms of the performance of the search engine, the artificial intelligence algorithms that recognize speech, and the text-to-speech engine, among other capabilities.

Eight participants agreed that Siri did not perform as effectively as Amazon Echo and Google Assistant in processing speech input and understanding complex commands, maintaining the context of conversation when consecutive questions were asked, and generating intuitive responses or feedback. For example, P7 found himself being able to input voice requests more naturally using Echo, while with Siri, he found that he had to consciously think about ways in which a query/command could be presented in a refined format so that it would be recognized on the first attempt.

P11 found that Google Assistant offered the most potential when attempting to perform follow-up queries based upon the response presented by the technology. He reported that he did not need to spend much time thinking about how to formulate the follow-up query, in contrast with other VAPAs: "*Siri has fallen way behind in the past couple of years. Google Assistant is ahead by a mile for follow-ups.*" Other participants explained that, compared to their experience with Amazon Echo or Google Home, Siri does not seem as responsive. P7 appreciated interactions with Echo, as it seemed more like speaking to another person, rather than a device. He said, "*There's a certain human quality to [Echo] that you don't experience especially with Siri. It will actually talk to you, whereas Siri talks at you.*"

While tasks could be performed more expediently using certain VAPAs, five participants noted frustration with making more complex queries with Siri, because it would provide a list of search results the user needed to sort through. In contrast, they would often receive more intuitive responses when asking complex questions of their Echo or Google Home. Users with both home and mobile VAPAs preferred using their home VAPAs when possible, as this was considered to be more reliable.

When discussing the merits of VAPAs, some participants considered the reputation/performance of the companies who make the product. For example, P8 indicated that the Google Search web interface provides targeted search results, and therefore Google Assistant would respond to queries more appropriately than other platforms. P5 similarly inferred that he knew Google Home would be a valuable purchase, as he had previously had positive experiences with Google products and thought highly of the organization. He attributed the ease of asking Google Home to do many things at the same time—such as playing a specific music playlist on his kitchen smart speaker, setting multiple timers, and asking to walk him through a cooking recipe step-by-step—to the quality of the Google Search website.

P2 appreciated the flexibility of controlling external devices, such as smart lights, through both Siri and Amazon Echo. It provided her reassurance that, should one VAPA fail or be unavailable, she could still control the technology using the other VAPA. In contrast, two iPhone users (P9 and P10) stated that they performed the majority of tasks using Siri (pre-installed on and part of iOS). While Google Assistant (a third-party app) was thought to offer a superior experience, Siri could be initialized much faster by simply pushing the home button on their mobile device.

Using Mobile VAPAs in Public

Seven participants preferred not to interact with Siri frequently while commuting by public transport due to concerns about disclosing sensitive identifying information, or having strangers overhear their personal affairs. For example, P2 emphasized that maintaining situational awareness should be considered as top priority when commuting. She would only use Siri to inquire about her current location and nearby points of interest.

To overcome interaction challenges when ambient sounds were present, participants in our study largely opted to use a headset when interacting with their VAPA on their mobile devices. The headset's microphone was thought to pick spoken information up more effectively compared to the built-in microphone on a smartphone. In addition, by using a headset, participants could stow their mobile device in their pocket or bag. This was thought to serve two purposes. Firstly, as the device would be out of view from third parties, there would be a lower likelihood of the device being stolen. Secondly, as the hands are often encumbered with a cane or service animal harness, holding a phone in the same hand while navigating may prove challenging.

Social Awkwardness and Disruption to Communication

Nine participants described the awkwardness associated with using mobile VAPAs in public. Concerns were expressed regarding drawing attention from others when voicing commands, as it may be considered socially-inappropriate by sighted others to interact with a mobile device at particular times. Furthermore, if onlookers could not see that the user was speaking into a mobile device, they may perceive the user to be talking to themselves. As a result, participants were found to limit their interactions with VAPAs in public (e.g., at social gatherings, meetings). P3 recalled an unpleasant experience in one of her classes at school where Siri on her iPad notified her to brush her teeth in a loud voice while in the presence of others, "*I calmed myself down by saying it was good that people saw that a regular iPad can do these things and that we don't need special pieces of equipment. But that's what I just told myself to feel better.*" P6 mentioned that she would not use Siri for texting in the presence of others including her friends because she felt self-conscious, "*I think texting blind is significantly more socially impeding than texting sighted. I wouldn't like to continue extended text conversations with Siri. It's awkward.*"

Using VAPAs at Home

In contrast to awkwardness in public settings, participants generally indicated their home space was considered to be a "comfort zone" where they would have no concerns interacting with VAPAs, even without headsets. P14 stated that she would interact with Siri with the volume of output increased at home, without the worry of disturbing others, or without looking out of place.

Four participants described the convenience associated with having their home VAPA perform a wide range of tasks such as controlling a thermostat or microwave. P2 found that VAPA interaction with a thermostat would allow her to plan for when she returned home. She recounted, "*So by the time I get home the heat will be on, and will turn off at the time I want. Routine is nice ….*" Participants highlighted that a greater sense of independence could be achieved, as the user themselves could control more devices within their living space. P8 and P9 considered their home VAPAs to be a "*perfect fit*", which had become an inseparable part of their lives. P8 used his Amazon Echo several times a day to manage his daily routines by setting multiple reminders or timers, "*So even I'll ask it to remind me to take out the rubbish on Thursdays. I can't imagine being without it at this point.*" VAPAs were described by P10, as helping to put blind and sighted individuals on an even playing field when performing tasks. For example, regarding using his VAPA to control game interfaces, he explained: "*as there's no screen [needed], we [blind and sighted users] are all able to interact on the same level when playing games*."

Privacy Concerns

The issue of privacy arose when discussing VAPA usage. While participants highlighted worries about using these technologies in public, in contrast, there appeared to be relatively few privacy concerns when using them at home in the presence of close friends or family members (e.g., when listening to notifications/reminders). Participants found sharing home-based VAPAs to be a beneficial feature that allows partners or family members with diverse abilities to take equal advantage of the services of an intelligent companion in the home setting. However, worries were expressed that emergence of new home VAPAs featuring a visual display like the Echo Show would encourage designers to move away from improving voice-based interaction, and focus on designing inaccessible content – namely, graphical information.

Participants were well aware of stories in the news relating to data breaches, and 'dolphin attacks' where VAPAs were thought to listen to conversations. Participants were generally willing to accept the threat of theft or misuse of data as a reasonable compromise for gaining access to perform tasks which may otherwise be challenging or cumbersome. Interestingly, multiple participants indicated that their personal data had already been exposed in several other ways (e.g., through their Gmail and iCloud accounts), so they would not be worried about further breaches from VAPAs. For example, P14 stated, "*Yeah sure… you don't use that experience at all, or you take advantage of the device to the fullest. You can't have your privacy and have it be such a convenient companion. To me, it's worth putting that information out there*."

DISCUSSION AND IMPLICATIONS FOR DESIGN

For years, assistive technology researchers have been studying the design of accessible voice-based interfaces for the blind community [2,14,18,19,21]. The mainstream adoption of VAPAs like Google Siri, Amazon Echo, Google Home, and Microsoft Cortana [8,10,11,16] presents perhaps one of the best opportunities yet for blind individuals to have access to an affordable, uniform, screenless interface to a variety of digital apps that does not immediately stigmatize. Notably, Pradhan, Mehta, and Findlater [21] have recently shown that a broad range of people with disabilities—and, prominently, people with vision impairments—have already embraced the Amazon Echo platform as a practical accessibility solution to many daily challenges. The present study extends this work by comparing accessibility challenges across VAPA platforms—Siri, Google Assistant, and Google Home in addition to Amazon Echo—not only in the home setting but also while in public or in presence of others.

While confirming previous findings that voice interaction is convenient for blind people [2] and VAPAs like Echo are one emerging platform in this community [21], our work reveals new categories of benefits and challenges to adoption. Voice interactions were found to limit time spent performing gestures and were in many cases preferred to mobile touchscreen interactions. Participants appreciated the ability to access an array of applications/services using just one VAPA interface. VAPAs provided alternative, accessible interactions with apps on their mobile (e.g., calendar) and off-the-shelf home technologies (e.g., thermostat). This removed the need to purchase expensive specialist technologies solely catered to people who are blind. While prior work documents that VAPAs are used by individuals who are sighted more for entertainment purposes [16], our study found that these technologies are serious tools for individuals who are blind, as they make possible day-to-day tasks which others may take for granted. In turn, VAPAs played an important role in evoking feelings of independence and empowerment among blind users.

This study additionally surfaced new technical challenges that need to be overcome. Some of these have been previously documented and affect users regardless of ability (e.g., misinterpretation of commands in certain situations like noisy environments). However, because blind users desire to accomplish more complex tasks through their VAPAs, issues such as identifying and recovering from recognition errors, difficulties with timeouts, presentation of auditory content at a slow speed, among others, disproportionately impact individuals who are blind. While sighted users are able to overcome these through alternative interaction modes (e.g., fixing dictation errors visually), lack of appropriate and equal alternatives for blind users motivate renewed attention to these usability issues.

Designing to Meet User Needs

Our findings offer an insight to the experiences of individuals who are blind interacting with VAPAs. To promote inclusivity, we suggest additional considerations should be made when designing VAPA interfaces.

Supporting Voice Interaction – Our study highlights the diversity of commands/queries made by individuals who are blind (Table 2). These range from simple to complex commands. Design should better support users to ensure that input errors are detected in a timely manner, enabling users to recover appropriately. VAPAs should better react to longer, more complex commands, as performing these using voice input offers a simpler alternative to input using traditional methods [2]. Solutions should also take into account that users have different needs and abilities, and be designed accordingly (i.e., more time to perform tasks prior to the system timing-out).

Providing Relevant Output and Control of Output – Findings from our study have shown that responses to queries presented may not provide the appropriate level of succinctness/detail needed, particularly in scenarios where the user is multi-tasking or on-the-go. Control should be afforded to allow users to customize the output, depending on preference and context of use. This may also speed-up the process of being able to detect errors made, and recover from these errors.

Synchronization between Visual and Non-Visual Cues – Voice interfaces were described as being "equitable" in nature – meaning that the user's level of sight should ideally not make a difference when interacting with the technology. However, lack of non-visual alternatives to graphical cues were identified (e.g., colored lights flashing on Amazon Echo). While speech-based equivalents of these cues may be useful, these may end up overloading the user. Short bursts of audio (earcons) could be used to bridge the gap between the visual and non-visual representation of notifications and alerts provided by VAPAs. These would not only provide support to individuals who are blind who would be able to detect these cues in a similar time frame to their sighted peers, but also support tasks which are collaborative in nature between blind and sighted users, both at home and while on-the-go.

Supporting Access to Third Party Devices - Findings highlighted that blind participants control other devices through their VAPAs. In certain cases, VAPAs may be the only form of technology that allows individuals who are blind to remain independent when devices are inaccessible (e.g., thermostats, light switch interfaces, and home appliances like microwave ovens with touch screen interfaces). Design should account for these voice-only interactions, and consider scenarios where multiple devices may be used to control the same technologies. As stressed by our participants, the trend of adding visual displays to home VAPAs (e.g., Amazon Echo Show) may lead to the creation of a gap between blind and sighted usage with these technologies. Future design of VAPA technology should consider equal visual and non-visual affordances of the interface to preserve the equality of access by users with or without sight.

Accounting for Privacy and Situational Factors – While other researchers have described issues relating to security and privacy (e.g., [1,4,9]) along with respective workarounds, issues of maintaining privacy from third parties, and worries regarding the social impact of using VAPAs in public settings, can influence user behavior. Design should ideally take into account that mobile VAPAs will be used in a range of environments and may need to respond appropriately depending on the situation. In certain contexts, providing "whisper" or "discreet" modes where personal details are either presented more quietly or are masked, may provide the privacy necessary for users.

Fostering Trust in VAPAs – Participants took additional time and effort to verify whether the commands which had been made using their VAPA (e.g., calendar entries) had indeed been created. Errors had led to challenges in trust in using VAPAs, leading to workarounds. Additional feedback could offer promise to individuals who are blind, once certain types of complex commands are performed, as this may help to alleviate concerns about whether these were performed correctly. Additionally, incorporation of appropriate feedback may better support "discoverability", an important factor highlighted in prior work [21].

Greater Understanding of User Needs - Participants in our study felt more comfortable using VAPAs in the privacy of their own home for performing both short and complex tasks. Providing additional support for these tasks may not only aid individuals who are blind, but may also play a role in supporting all users regardless of ability who may be distracted. To gain a greater understanding of the needs of individuals who are blind, participatory design with target populations may offer one method of helping to design VAPA technologies which match user needs and abilities.

LIMITATIONS OF THE STUDY

While the sample size selected for the study was in line with other qualitative studies examining the needs of individuals who are blind (e.g., [1,4]), it was difficult to identify differences in experiences and behaviors by status of sight (i.e., congenitally vs adventitiously blind). Participants who were recruited were typically young in age and proficient with technology. It would be interesting to compare and contrast experiences of groups with different age ranges and technology expertise, particularly in regards not only to use of the VAPAs but also other factors which should be considered when evaluating experiences (e.g., the ability to set-up VAPA devices/technologies).

CONCLUSION AND FUTURE WORK

In this paper, we described a study investigating blind users' experiences with VAPAs. Providing increased levels of control, designing for privacy and situational factors, examining ways to foster trust, and synchronizing presentation of visual and non-visual cues were thought to improve the quality of the user experience. The next logical step in the research is to examine sighted users' attitudes towards VAPAs, along with details of their experiences using these technologies. This would provide a more detailed point of comparison. Observational studies will be conducted to provide a deeper insight into usage behaviors.

ACKNOWLEDGMENTS

We thank Cary Chin and Sidas Saulynas for their help with data gathering and analysis. We thank our participants for sharing their time and expertise.

REFERENCES

1. Ali Abdolrahmani, Ravi Kuber, and Amy Hurst. 2016. An empirical investigation of the situationally-induced impairments experienced by blind mobile device users. In *Proceedings of the 13th Web for All Conference (W4A'16)*, 21:1–21:8. https://doi.org/10.1145/2899475.2899482
2. Shiri Azenkot and Nicole B. Lee. 2013. Exploring the use of speech input by blind people on mobile devices. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'13)*, 11:1–11:8. https://doi.org/10.1145/2513383.2513440
3. Emily C. Bouck, Sara Flanagan, Gauri S. Joshi, Waseem Sheikh, and Dave Schleppenbach. 2011. Speaking math — a voice input, speech output calculator for students with visual impairments. *Journal of Special Education Technology* 26, 4: 1–14. https://doi.org/10.1177/016264341102600401
4. Stacy M. Branham, Ali Abdolrahmani, William Easley, Morgan Scheuerman, Erick Ronquillo, and Amy Hurst. 2017. "Is someone there? do they have a gun": how visual information about others can improve personal safety management for blind individuals. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'17)*, 260–269. https://doi.org/10.1145/3132525.3132534
5. Stacy M. Branham and Shaun K. Kane. 2015. The invisible work of accessibility: how blind employees manage accessibility in mixed-ability workplaces. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS'15)*, 163–171. https://doi.org/10.1145/2700648.2809864
6. Mei-Ling Chen and Hao-Chuan Wang. 2018. How personal experience and technical knowledge affect using conversational agents. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion (IUI'18)*, 53:1–53:2. https://doi.org/10.1145/3180308.3180362
7. Abide Coskun-Setirek and Sona Mardikyan. 2017. Understanding the adoption of voice activated personal assistants. *International Journal of E-Services and Mobile Applications (IJESMA)* 9, 3: 1–21. https://doi.org/10.4018/IJESMA.2017070101
8. Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What can i help you with?": infrequent users' experiences of intelligent personal assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'17)*, 43:1–43:12. https://doi.org/10.1145/3098279.3098539
9. Aarthi Easwara Moorthy and Kim-Phuong L. Vu. 2015. Privacy concerns for use of voice activated

personal assistant in the public space. *International Journal of Human-Computer Interaction* 31, 4: 307–335. https://doi.org/10.1080/10447318.2014.986642

10. Christos Efthymiou and Martin Halvey. 2016. Evaluating the social acceptability of voice based smartwatch search. In *Information Retrieval Technology* (Lecture Notes in Computer Science), 267–278. https://doi.org/10.1007/978-3-319-48051-0_20
11. Ido Guy. 2016. Searching by talking: analysis of voice queries on mobile web search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'16)*, 35–44. https://doi.org/10.1145/2911451.2911525
12. Sangyeal Han and Heetae Yang. 2018. Understanding adoption of intelligent personal assistants: a parasocial relationship perspective. *Industrial Management & Data Systems* 118, 3: 618–636. https://doi.org/10.1108/IMDS-05-2017-0214
13. A. Helal, S. E. Moore, and B. Ramachandran. 2001. Drishti: an integrated navigation system for visually impaired and disabled. In *Proceedings Fifth International Symposium on Wearable Computers*, 149–156. https://doi.org/10.1109/ISWC.2001.962119
14. Jonathan Lazar, Aaron Allen, Jason Kleinman, and Chris Malarkey. 2007. What frustrates screen reader users on the web: a study of 100 blind users. *International Journal of Human–Computer Interaction* 22, 3: 247–269. https://doi.org/10.1080/10447310709336964
15. Irene Lopatovska, Katrina Rink, Ian Knight, Kieran Raines, Kevin Cosenza, Harriet Williams, Perachya Sorsche, David Hirsch, Qi Li, and Adrianna Martinez. 2018. Talk to me: Exploring user interactions with the Amazon Alexa. *Journal of Librarianship and Information Science*: 0961000618759414. https://doi.org/10.1177/0961000618759414
16. Ewa Luger and Abigail Sellen. 2016. "Like having a really bad PA": The gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'16)*, 5286–5297. https://doi.org/10.1145/2858036.2858288
17. Niina Mallat, Virpi Tuunainen, and Kristina Wittkowski. 2017. Voice activated personal assistants–consumer use contexts and usage behavior. In *Technology Research, Education, and Opinion (TREO), Americas Conference on Information Systems (AMCIS'17)*.
18. Austin M. Mulloy, Cindy Gevarter, Megan Hopkins, Kevin S. Sutherland, and Sathiyaprakash T. Ramdoss. 2014. Assistive technology for students with visual impairments and blindness. In *Assistive Technologies for People with Diverse Abilities*. Springer, 113–156. https://doi.org/10.1007/978-1-4899-8029-8_5
19. Emma Murphy, Ravi Kuber, Graham McAllister, Philip Strain, and Wai Yu. 2008. An empirical investigation into the difficulties experienced by visually impaired Internet users. *Universal Access in the Information Society* 7, 1–2: 79–91. https://doi.org/10.1007/s10209-007-0098-420.
20. François Portet, Michel Vacher, Caroline Golanski, Camille Roux, and Brigitte Meillon. 2013. Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects. *Personal and Ubiquitous Computing* 17, 1: 127–144. https://doi.org/10.1007/s00779-011-0470-5
21. Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. "Accessibility came by accident": use of voice-controlled intelligent personal assistants by people with disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, 459:1–459:13. https://doi.org/10.1145/3173574.3174033
22. Gianluca Schiavo, Ornella Mich, Michela Ferron, and Nadia Mana. 2017. Mobile Multimodal Interaction for Older and Younger Users: Exploring Differences and Similarities. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM'17)*, 407–414. https://doi.org/10.1145/3152832.3156615
23. Waseem Sheikh, Dave Schleppenbach, and Dennis Leas. 2018. MathSpeak: a non-ambiguous language for audio rendering of MathML. *International Journal of Learning Technology* 13, 1: 3–25. https://doi.org/10.1504/IJLT.2018.091609
24. Derrick W. Smith and Stacy M. Kelly. 2014. Assistive technology for students with visual impairments: A research agenda. In *International Review of Research in Developmental Disabilities*. Elsevier, 23–53. https://doi.org/10.1016/B978-0-12-420039-5.00003-4
25. Linda Wulf, Markus Garschall, Julia Himmelsbach, and Manfred Tscheligi. 2014. Hands free - care free: elderly people taking advantage of speech-only interaction. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI'14)*, 203–206. https://doi.org/10.1145/2639189.2639251
26. Yu Zhong, T.V. Raman, Casey Burkhardt, Fadi Biadsy, and Jeffrey P. Bigham. 2014. JustSpeak: enabling universal voice control on Android. In *Proceedings of Web for All Conference (W4A'14)*, 36:1–36:4. https://doi.org/10.1145/2596695.2596720