Exploring Smart Speakers as Cognitive Tracking Tools

Abstract. Dementia has no cure but some symptoms can be treated. Therefore, early detection could greatly improve a person’s quality of life. Smart speakers offer a potential platform for voice user interface (VUI) tools that can aid the diagnosis of cognitive decline and dementia in older adults. A cognitive screener was implemented as a VUI with an accompanying mockup of a clinical dashboard. Experts were interviewed about their perceptions of both artifacts. We evaluated our prototype with seniors via a lab-based usability study and a 5-week field deployment. Results show the approach is promising and reveal various design implications for future fully implemented smart speaker based cognitive tracking systems.

Keywords. Voice User Interface; conversational user interface; voice based user interface; cognition; tracking; evaluation; System Usability Scale, Callahan Six-Item Screener; older adults; elderly; aging; elder care

1 Introduction

By 2060 there will be about 98 million Americans over the age of 65 and as more people live longer, a larger percentage of the population will be effected by cognitive decline [1]. Dementia is not one disease; rather it is a label that encompasses multiple symptoms that reduce a person’s ability to complete daily functions. Currently, there is no cure or known treatment that can stop its progression [2]. However, some symptoms and diseases that co-occur with it can be treated [3]. Therefore, early detection of dementia could greatly improve a person’s quality of life. This makes dementia diagnosis tools vital [4].

Unfortunately, many current techniques for diagnosing dementia require special equipment, are time consuming, or costly. This may be why the Alzheimer’s Association found that doctors do not routinely test older adults’ memory and cognition, even though most adults thought such testing is important [5]. This pattern suggests doctors may need simpler and quicker alternative diagnosis techniques to meet older adults’ desires for regular cognition checks.

Almost one in five American adults use a smart speaker in their home to interact with a virtual assistant and older adults have expressed interest in such devices [10]. Currently 10.1 million older adults (53
years and older) use voice-based assistants [11]. Companies have explored using smart speakers and other platforms to provide healthcare solutions for this population, but there is a lack of research investigating how older adults interact with their smart speaker devices [12][13].

This study makes 3 contributions to the Interact community. First we adopted a cognitive screener into a VUI. Second, we interviewed experts about the utility of such an approach. Finally, we evaluated our prototype with seniors via a lab-based usability study and a 5-week deployment.

2 Related Work

2.1 Clinical Cognitive Screeners

There are numerous short dementia screening tools. The Mini-Mental State Examination (MMSE) is a tool for detecting warning signs of dementia that can be administered in about 10 minutes [6-7]. Although the MMSE is the most popular screener, some researchers have raised doubts about its capacity for detecting mild cognitive impairment (MCI). MCI is the stage between healthy aging and dementia. Such concerns led to the development of other screeners such as the MoCA. The MoCA has been found to be a better detector of MCI than the MMSE in adults 60 and older [8]. However a limitation of the MoCA and MMSE is that they require the use of paper and writing materials. A different test, the Callahan Six-Item Screener, has been validated in multiple countries and performs similarly and sometimes better than the MMSE at detecting cognitive decline [14][8][9]. This test can be administered in minutes and does not require any physical supplies.

2.2 Assessing Cognition with Interactive Systems

Researchers have attempted to measure older adults’ cognitive statuses using a variety of technology platforms. For example, Nonaka et al. explored presenting a computer based audio interactive agent to participants with dementia [15]. Results revealed that certain characteristics of speech were correlated to cognitive status. A game based approach was used by [16]. The participants’ solitaire games were tracked over time to watch for performance changes that could indicate general cognitive decline. Sensors have also been used in home environments
with the goal collecting behavior patterns indicative of cognitive decline [17]. Recently wearables [18] and mobile apps [19] [20] have also been used to identify cognitive decline. While the results of these studies are promising, the requirement of specialized software and hardware limits the chance they could become widely adopted.

2.3 VUIs for Older Adults

Researchers have started exploring older adults’ perceptions of VUIs. [21] found that older adults enjoyed using a mobile speech-only interface, but noted that participants’ impressions of the system decreased based on the quality of the interaction. Additionally, [22] found voice input was easy to learn to use, but could be challenging because of the short time allowed for issuing commands. Relatedly, [23] discovered mild cognitively impaired older adults could use voice controlled virtual agents. Older adults also provided positive feedback when using voice to control smart home systems [24-26].

VUIs are also garnering attention by private companies [27]. One retirement home study found that Amazon Alexa improved seniors’ sense of wellness [28] and another found that Echo Dots could be used to gather information about seniors living with diabetes [29]. Another study found that reporting behaviors via a VUI led to increased sense of accountability and compliance [30].

This study extends the current understanding of older adults’ interactions with VUIs and cognitive tracking tools by exploring the feasibility of utilizing smart speakers as a new modality for cognitive test delivery.

3 Prototype Design and Implementation

The Callahan Six-Item Screener [9] was implemented as an action using the Google Dialogue Flow developer interface. The action was named “Weekly Check In” (WCI) to remove possible negative connotations of being a cognitive assessment tool. There were a few changes made to [9] during its conversion to an action. First, the three-word recall portion was adopted to avoid learning effects and boredom during the weekly administration. Three words were selected at random from a list of items compiled from the Mini-Cog [31]. The action start-
ed with the verbal command: “Ok Google, talk to Weekly Check In.” When the WCI was completed it would say, “Thank you.”

The WCI was programmed to provide a series of responses if it did not recognize the participant’s input. This included repeating the question and encouraging a response by using helper-phrases (e.g., “Please repeat that” or “Please try again.”). If the participant’s input was not recognized after two helper-prompts, the action would provide an apology and end the interaction.

Axure design software was used to create the dashboard wireframes. It included information about scores, performance over time, user inputs, time of activation, and the option to view exact transcripts of all interactions with the WCI. The dashboard was not interactive.

4 Subject Matter Expert Interviews

4.1 Methods

The university’s ethics committee approved the study protocol and all participants provided informed consent before participating in the study. All participants were from a southern American city. Recruitment was conducted over the phone or by email and participants received no compensation. (This includes participants in 5.1 and 6.1).

Participants. We met with seven subject matter experts and cognition and aging specialists serving as professors, university researchers, or staff researchers from organizations focused on aging research. Experts had knowledge of cognitive science, geriatrics, or adult care.

Procedure. Semi-structured interviews gathered information about cognitive tracking tests, the diagnosis of mental decline, and impressions about the WCI and results dashboard mockup. Questions were designed to learn about specific cognitive decline tools and testing procedures, how cognitive test results are communicated, and issues with current cognitive testing strategies. Additionally, participants were prompted to provide thoughts about a demo of a person interacting with the WCI. All responses were compiled and grouped by question. Then the thematic analysis technique outlined in [32] was applied to the data.
4.2 Results

All participants believed a smart speaker administered screener would be potentially useful. There were four emergent themes and sub-themes from the interviews. First, experts mentioned verbally administered screeners will have limited scope and only test specific cognitive components. Therefore, a cognitive tracking tool would only be useful if it provided information about the abilities the expert was interested in testing. Second, there were concerns about the administration of the screener because participants could possibly cheat or have variable performance due to environmental conditions. Specifically, participants expressed concern with the smart speaker’s volume and error handling capabilities as well as possible practice effects that could occur from repeated use of the screener. The third theme focused on concerns regarding the communication of sensitive and personal results. Participants were particularly worried about how the data would be shared and doubted users would willingly provide data that could compromise their independent living. The experts wanted as much information about the results as possible, but speculated that primary users of the system would just want an indication of their cognitive status. The fourth emergent theme was the importance of tracking older adults over time. Participants mentioned that establishing a baseline would be essential and they would only want to be warned when a user significantly differed from that baseline.

5 Evaluation of the Weekly Check In (WCI)

5.1 Method

Participants. Four older adults over the age of 65 participated in this study. None of the participants (in this study or 6.1) owned a smart speaker device. Additionally, no participants had self reported cognitive impairments, mental disorders, or hearing loss that would interfere with the use of a VUI.

Procedure. Participants came to a university lab to complete this assessment. First participants were introduced to the concept of using Google Home and Google actions because they had no previous experience with launching one. Then the participants were told to try to start the WCI. No additional information about the action’s purpose or fea-
tures were provided. All participants were able to successfully start the action; however, they all expressed concern about being able to say the correct order of words to launch it in the future. Once the participants completed interacting with the WCI action, they answered a series of semi-structured interview questions focused on the usability of the action, their results reporting preferences, and the extent the action tested their cognition.

5.2 Results

The older adults’ interactions with the WCI revealed a mix of qualitative feedback. Participants felt the action was easy to use and some thought it was like a game. Participants wanted immediate feedback after using WCI either a verbally presented score or some type of visual display. Participants were afraid they would forget how to launch the WCI and would need reminders to interact with it. Two of the participants struggled with speaking in a manner the Google Home could recognize. Once the participants repeated their commands, the speaker accepted their responses.

6 Field Deployment

6.1 Methods

Participant. A 65 year-old woman took part in the study. Procedure. This was a 5-week in home usability study. The participant was told she could call the researcher for technical support as needed. During the first session, the experimenter a) deployed the Google Home and connected it to her wireless home network b) conducted the usability study (see below) and c) gathered the participant’s initial impressions of interacting with the VUI. The participant was instructed to use the Google Home however she desired.

One week after the first session, the researcher conducted a four question semi-structured interview. The experimenter then introduced the WCI by giving a brief description of its function. Next, the participant tried it once. The participant was instructed to use the WCI at the same day and time for the next three weeks. The participant accepted the offer to have a weekly reminder. The experimenter instructed the participant to continue using the Google Home however she desired.
During the next three weeks the experimenter used a Google web portal to ensure the participant used the smart speaker daily.

On the final day of the study (week 5), the experimenter conducted a semi-structured interview to learn the participant’s feelings about the Google Home, the WCI action, and any difficulties that were encountered. Lastly, the Google Home was deactivated from the participant’s network, reset, and removed.

**Usability Test.** The usability test followed a procedure similar to [33]. The researcher explained the basic technique for interacting with the Google Home, then present the participant with a list of tasks including: play some music, reduce the volume, inquire details on the song being played, stop the music, set an alarm for the following day at 6 a.m., start a timer for 30 seconds, inquire about the current weather conditions, learn about Steve Jobs, hear a joke, determine the time, and calculate 100 divided by 10. The participant was instructed to complete the tasks sequentially and verbally indicate when each task was finished. A stopwatch was used to record task completion tasks. Once the 11 tasks were done the participant fill out a paper version of the SUS.

### 6.2 Results

The SUS scores over weeks 1, 3, and 5 were respectively: 82.5, 100, and 95. The user took 155.01 seconds to complete the usability tasks in the first week. The qualitative data revealed three major themes. Privacy fears were the most prevalent topic. Those fears were exacerbated by online and TV news related to privacy leaks and home systems “listening all the time.” Second, the Google Home was seen as a fun, but unnecessary accessory. After the first week the participant felt mastery over the basic functions and was not motivated or sure how to independently try new functions. Finally, the VUI platform presented challenges including error handling, speech recognition, and speech speed. Additionally, the participant was unsure when she provided invalid inputs compared to when the speaker had a system level problem (e.g. no Wi-Fi connection.)

### 7 Discussion

Our study focused on exploring the feasibility of using a smart speaker based cognitive screener, Weekly Check In (WCI), as a way to
track cognition in older adults. Clinical experts were cautiously optimis-
tistic that our approach could eventually provide valid and reliable
patient data. All seniors in our usability study were able to complete the
WCI tasks. Finally, participants considered the WCI entertaining and
wanted immediate feedback about their performance. We also conduct-
ed a 5-week deployment. The participant had an overall positive expe-
rience with the Google Home and WCI. The SUS ratings she gave indic-
ating the interface was well received, but interviews revealed lingering
concerns about privacy and independent technical problem solving.
Taken together, these results show an in-home dementia screener could
be useful and usable to both physicians and seniors.

Our study revealed numerous design implications. Experts stressed
that visualizing long-term results would be important for clinicians.
Increasing the discoverability of new features could lead to sustained
smart speaker use among seniors. Likewise, automated reminders of the
verbal commands and easy to follow error handling would increase
compliance and usability of systems such as WCI. The in-home de-
ployment participant’s perception that Google Home was superfluous
shows that individuals who do not already own this device, mobile or
web-based VUIs might be a better option.

This study has a number of limitations that should be addressed in
future research. First, it was conducted in the U.S. and seniors in other
parts of the world may have different concerns. Our mockup was lim-
ited, so next steps should include more exploration of visual data re-
porting options for both older adults and doctors. Finally, we picked [9]
because it had been translated into various languages, but it could be
that other screeners are equally useful.

8 Conclusion

Our approach leveraged the power of short verbal screeners [14][8-9]
and seniors’ ability and willingness to use VUIs [22] in their health and
wellness routines [12-13]. Our study also shows that a simple and ac-
cessible cognitive tracking tool could be helpful in bridging the needs
of both physicians and older adults [5]. It provides compelling evidence
that this is an area rich for exploration.
9 References


