

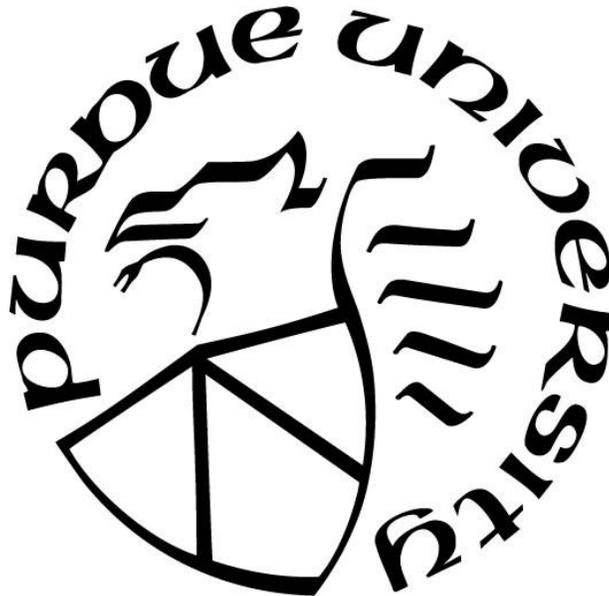
**A STUDY ON USERS' DISCOVERY PROCESS OF AMAZON ECHO
DOT'S CAPABILITIES**

by
Lingyi Zhang

A Thesis

*Submitted to the Faculty of Purdue University
In Partial Fulfillment of the Requirements for the degree of*

Master of Science



Department of Computer Graphics Technology

West Lafayette, Indiana

May 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Paul Parsons, Chair

Department of Computer Graphics Technology

Dr. Austin L Toombs

Department of Computer Graphics Technology

Dr. Mihaela Vorvoreanu

Department of Computer Graphics Technology

Approved by:

Dr. Nicoletta Adamo-Villani

Graduate Program Co-Chair

Dr. Colin Gray

Graduate Program Co-Chair

ACKNOWLEDGMENTS

I would like to thank my committee for helping me refine the thesis topic, giving me suggestions on methods and all the other supports. It has been through a long way since my very first research question. I could not complete this thesis without their help. Thank you.

TABLE OF CONTENTS

LIST OF TABLES	VI
LIST OF FIGURES	VII
ABSTRACT.....	VIII
CHAPTER 1. INTRODUCTION	1
1.1 Statement of Problem.....	1
1.2 Research Goal	2
1.3 Research Questions	2
1.4 Significance.....	3
1.5 Assumptions.....	4
1.6 Limitations	4
CHAPTER 2. Review of literature.....	6
2.1 History of Virtual Assistants.....	6
2.2 Previous Studies on Virtual Assistants	11
2.2.1 Evaluating and Comparing Virtual Assistants	11
2.2.2 Virtual Assistants in specific using scenarios	13
2.2.3 Current challenges	15
2.2.4 Summary	17
CHAPTER 3. METHODOLOGY	18
3.1 Process	18
3.2 Sampling	19
3.3 Data Collection	19
Dialogue History and	19
3.3.1 Dialogue History and Surveys	19
3.3.2. Follow-up Interview.....	20
3.4 Data Sources	21
3.5 Analysis.....	21
3.6 Trustworthiness.....	23
3.7 Research Identity Memo.....	24
3.8 Summary	25
CHAPTER 4. Results.....	26
4.1 Overview.....	26
4.1.1 Sample.....	26
4.1.2 How are the results related to research questions	27

4.2	Theme 1: The Activities Participant Engaged to Learn the Capabilities of Echo Dot .	28
4.2.1	“Alexa, how’s the weather?”	28
4.2.2	Exploring with the Alexa App	30
4.2.3	Google It	33
4.2.4	Ask Alexa or Ask a Friend?.....	35
4.3	Theme 2: The Influence of Echo Dot Being Screen-less.....	38
4.3.1	Talking to Echo Dot is not Easy	38
4.3.2	Where is the confirmation.....	40
4.3.3	Privacy and Security	42
4.4	Factors of Participants that Influenced the Initial Use	42
4.4.1	Language Barrier	42
4.4.2	Participant’s Familiarity of Technology and Their Motivations	43
4.4.3	Previous Experience with Virtual Assistant	45
4.4.4	Participants’ views of Echo Dot	46
4.4.5	The Decision on the Placement of Echo Dot	48
CHAPTER 5. Discussion and Conclusion		51
5.1	Discussion.....	51
5.1.1	The Proper Help that Echo Dot and Alexa App Should Provide.....	51
5.1.2	Timely and Accurate Feedback for the Users.....	55
5.1.3	Building Proper User Expectations.....	56
5.1.4	Privacy	58
5.2	Limitations	59
5.3	Conclusion	61
APPENDIX A. Questions in THE SURVEYS.....		62
APPENDIX B. Interview Protocol		63
REFERENCES		65

LIST OF TABLES

Table 1: Participant Demographics.....	26
--	----

LIST OF FIGURES

Figure 1: Initial Codes	23
Figure 2: Alexa App.....	53

ABSTRACT

Author: Zhang, Lingyi MS
Institution: Purdue University
Degree Received: May 2019
Title: A Study On Users' Discovery Process Of Amazon Echo Dot's Capabilities
Committee Chair: Paul Parsons

With Home Virtual Assistant (HVA) becoming popular, it has entered many families. This study aims to understand how novice users explore the capabilities of HVAs and the factors influence the users' discovery process. The researcher used a study inspired by diary study, combining survey and interview to collect the data, along with dialogue history that contains all the activities between the user and the device. The results reveal the activities the users engage in and the factors of the HVA device and the users that influence the discovery process. Three themes emerged from the data, they are: the activities participant engaged to learn the capabilities of Echo Dot, the influence of Echo Dot being screen-less and factors of participants that influenced the initial use. Based on the results, the researcher discusses the current issues of the Echo Dot's design and suggests potential improvement for the HVAs.

CHAPTER 1. INTRODUCTION

This chapter introduces the statement of problem, research goal, research questions, significance, assumptions and limitations of this study.

1.1 Statement of Problem

With the recent development of machine learning, conversational agents have become significant components in major operating systems. Apple's Siri, launched in 2011, has been described as impressive and innovative because of its combination of voice recognition and natural language processing (Aron, 2011). Siri listens to users' orders, makes sense of the orders, and reacts to them. Siri has exposed the imagination of natural language processing technology to the wide public (Moore, 2016). It is available on all the Apple's major devices including Apple TV and HomePod (Apple, n.d.), which means users can access Siri from everywhere in their daily lives. Siri has brought the considerable possibilities not only in commercial use but also in the field of academia.

Following Apple, Google launched Google Assistant in 2012 (Google, n.d.), Microsoft launched Cortana in 2014 (Microsoft, n.d.), and Amazon launched Alexa in 2015 (Amazon, n.d.). Unlike Siri, existing in devices with screens, Alexa is a virtual assistant embedded in a speaker called Echo. Echo doesn't have a screen, it only replies to users with voice-based responses (Etherington, 2015). The release of Alexa and Echo introduced a new kind of device, so called smart speaker, or home virtual assistant (HVA). HVAs have changed the market. By August of 2018, the number of HVA devices starts to hit 50 million (Koetsier, 2018). And the market is still increasing quickly, approximately 48% annually (Koetsier, 2018).

It is believed that conversational interaction can improve the user experience (Vassallo, Pilato, Augello, & Gaglio, 2010). By entering users' daily routines and helping users completing tasks, HVAs have changed the ways of interacting with personal assistants. But current HVAs do not provide true natural language interaction (Cohen, Cheyer, Horvitz, El Kaliouby, & Whittaker, 2016), they require the users to learn its capabilities (Furqan et al., 2017). The user's initial use of HVAs is important as it defines how users use the devices in the future.

Voice-based interaction in users' daily life is still in lack of empirical research (Luger & Sellen, 2016; Porcheron, Fischer, Reeves, & Sharples, 2018; Purington, Taft, Sannon, Bazarova, & Taylor, 2017), even fewer studies have looked into how novice users discover HVAs' capabilities.

1.2 Research Goal

The goal of this study is to understand what is novice users' discovery process of the HVAs' capabilities. The study aims at exploring the activities users conduct to learn the HVAs' capabilities, how does the fact of being screen-less influence users' discovery and what difficulties do users encounter with HVAs. The ultimate goal is to provide design guidelines for HVA designers to help users discovering HVAs' capabilities by identifying users' activities, influencing factors of HVA and users in the discovering process. Amazon Echo Dot will be studied in this research, which helps the researcher to narrow the focus and be more specific about the context.

1.3 Research Questions

This research has one guiding research question as "what is the users' discovery process of HVAs' capabilities?". While this study will focus on Amazon Echo Dot because the price of

Amazon Echo Dot is relatively low compares to other HVAs and the funding of this study is limited. Thus this research has one research question and three sub research questions:

What is the users' discovery process of Amazon Echo Dot's capabilities?

- I. What activities do users engage in to discover Amazon Echo Dot's capabilities?
- II. How does Amazon Echo Dot being screen-less influence users' discovery process?
- III. What factors of the users influence their initial use of Echo Dot?

1.4 Significance

Previous studies have focused on social roles of virtual assistants (Folstad et al., 2018; Portela & Granell-canut, 2017; Purington et al., 2017), how people interact with virtual assistants (Corbett & Weber, 2016; Cowan et al., 2017; Porcheron, Fischer, & Sharples, 2017) and evaluation of virtual assistants (Assefi, Liu, Wittie, & Izurieta, 2015; Portet, Vacher, Golanski, Roux, & Meillon, 2013). As a relatively new area, still little is known about the details of users' interaction with HVAs (Porcheron et al., 2018).

The market of HVA is growing rapidly. It is reported that about 16% of adults of the US own Amazon Alexa or Google Home. With the API of Alexa opening to third party developers, Alexa experienced rapid growth on its skills (Kinsella, 2017). The skills of Alexa drive the engagement and repeat use of users (Kinsella, 2018). With 300% growth of Alexa skill count in 2017 (Kinsella, 2017), Alexa skill count has surpassed 30,000 in the US (Kinsella, 2018). How users react to such enormous number of skills and how novice users discover those skills are still unclear.

This study is significant because it will provide understanding of how users discover the capabilities of Amazon Echo Dot from different perspectives, which has barely been explored by other researchers. This study is unique because it looks at novice users' discovery process

closely via surveys, dialogue history and interviews, which can potentially reveal the true discovery process in home setting.

The results of this study will have the potential to provide a deeper understanding at what does novice users' discovering process of Amazon Echo Dot's capabilities look like. This study could also produce insights that are beneficial for future HVA design. The findings of this study could bring design implications of helping users to learn the capabilities of HVAs.

1.5 Assumptions

The following assumptions were made in this study:

1. Participants will be honest with their experience on HVAs.
2. The surveys and follow-up interviews can reveal how users truly discover the capabilities of Amazon Echo Dot.

1.6 Limitations

This study has following limitations:

1. Ideally, a large sample size might reflect the true population and provides wider generalizability, but due to the time restriction and money, sample size will be limited to ten.
2. Ideally, all the HVA devices should be studied, as each device has different designs and functionalities. The users might have different discovery process. But due to limited funding of this study, Amazon Echo Dot will be used for its low price.
3. The users from different geographic locations could use different methods to explore the capabilities of HVAs. In this study, participants will only be recruited at Lafayette, IN because of the time restriction.

4. The age group will be limited to 18 to 70.

CHAPTER 2. REVIEW OF LITERATURE

2.1 History of Virtual Assistants

Virtual assistants have been an essential part in many smart devices like phones and tablets. Since the launch of Siri in iPhone (Aron, 2011), virtual assistants have been placed great expectations on. The naturalness of voice interaction was seen as an advantage comparing to Graphic User Interface (Rosenfeld, Olsen, & Rudnicky, 2001). The wide-spread use of virtual assistants opened the possibilities of completing tasks, asking questions and navigating through applications by just talking to the phone (Moore, 2016). Soon Microsoft, Amazon and Google followed, releasing products aiming to grab the market share.

The concept of virtual assistants can be traced back to last century. In the movie “2001: A space odyssey” (Thurman, 2008), people imagined talking to the artificial intelligence HAL. The imagination of talking to a machine has existed since 1960s. The first generation of conversational agents was born in 1966, ELIZA (Weizenbaum, 1966). ELIZA was given a set of rules, which contained in a script. Users can type sentences or statements on a typewriter to ELIZA, ELIZA then identifies the most important keywords and responds with typed sentences. The rules in the script are limited, only keywords existing in the script can trigger the response of ELIZA (Weizenbaum, 1966). Of course ELIZA does not have the ability to truly understand the sentences from users, but using rule-based small talk has been accepted and followed by other researchers (Kopp, Gesellensetter, Krämer, & Wachsmuth, 2005).

Soon systems like SHRDLU (Winograd, 1972) and GUS (Bobrow, Kaplan, Kay, & Winograd, 1977) emerged. Both SHRDLU and GUS were designed with systematic grammar. The motivation of developing such systems was to apply linguistic knowledge to computer programming. GUS, for example, was given a knowledge structure that allows it to deconstruct

sentences from users and better understand the sentences. GUS aims at serving at the area of travelling, which reduces the amount of data it requires. Though GUS cannot trick people as if it is a real person, it "... represents a beginning step towards the construction of an intelligent language understanding system" (Bobrow et al., 1977).

One of the earliest Spoken Dialogue System developed in the USA was Air Travel Information System (ATIS) (Hemphill, Godfrey, & Doddington, 1990). ATIS has a limited database, including information like flights, airplanes, airports etc. What's innovative about the creation of ATIS system was that to collect real utterances in the real world, the founders of ATIS used simulated environment, asking subjects to speak, collecting the utterances and transferring the utterances to corpus that would be used in ATIS system. Another contribution of ATIS was the standard process of evaluating corpus, which become an important aspect of future spoken dialogue system (Mctear & Callejas, 2016, p.55). Following researches continued throughout the 1990s, that was also the time researches on Voice-user Interface began. Voice-user Interface shares the same technology with Spoken Dialogue System, the differences lie on the purposes of different communities (Mctear & Callejas, 2016). Spoken Dialogue Systems are created by researchers and universities and aim at making contributions to the field. Voice-user Interfaces, on the other hand, are made by developers and companies for commercial use mainly.

The early voice-user interface systems were designed to assist self-service tasks via calling as those tasks had routines and limited variations (Wilpon, Acoustics, & Vol, 1990). Through phone call, the systems only detected isolated words (Rabiner & Levinson, 1990; Wilpon, DeMarco, & Mikkilineni, 1988), such as saying "collect" for a collect call. While the actual results of such systems were not acceptable because many users did not only speak the defined words in isolation. To make voice-user interface more accessible to generable public,

Wilpon et al. (1990) promoted the algorithm of speech recognition and their model showed the potential to capture the vocabularies in unconstrained fluent speech. Soon in 1997, Gorin, Riccardi, & Wright (1997) further improved the algorithm of speech recognition and freed the caller from permissible vocabulary. Their model does not put the effort on fully understanding what the caller says, instead, the model tries to capture the salient fragments, which largely increased the recognition accuracy. Their algorithms enable the system to identify call types through user's speech without restricting the user to use defined words.

The research of spoken dialog systems continued and researchers (Allen et al., 1995; Ferguson & Allen, 1998) started to make the systems more intelligent and give the systems the capabilities to solve complex problems. The Trains system (Allen et al., 1995), for example, a conversational agent built with spoken dialogue system and plan reasoning capabilities, could communicate with human agents and collaboratively define and execute plans to manage a railway transportation system. Building on the Trains system, Trips planning system was smarter and even able to solve problems by its own (Ferguson & Allen, 1998).

Along the journey of making spoken dialogue systems more intelligent, many new theories like information state approach (Traum & Larsson, 2003) emerged. Information state approach helps researchers to manage dialogues, reuse and update components in the system. New programming language also promoted the conversational agents. One of the most famous conversational agents created after 21th century, Alice chatbot system, was built with Artificial Intelligence Mark-up Language (AIML) (Shawar & Atwell, 2007). AIML has different categories of contents and each category has a rule to match the input from users. AIML helped building chatbots in many other domains like education, online business, entertainment and information retrieval (Shawar & Atwell, 2007).

Another line of researches and developments started using statistical and data-driven approaches to understand natural language. But this method requires huge amount of data to train and costs a lot more than rule-based model (Mctear & Callejas, 2016). No successful model has been built so far (Bellegarda, 2013). On the other hand, rule-based model kept improving and gave rise to framework like Active that aims at providing "... a toolset and an associated methodology that lowered the bar for creating intelligent applications" (Guzzoni, 2008). Active framework has a rational network building with concepts. The concepts are connected by relationships. This network provides an environment that can react to the input from users, searches the associated concepts throughout the network then gives commands to the system. The most successful application of Active framework is Apple's Siri (Bellegarda, 2013).

Siri was originally developed by Siri, Inc. Apple acquired Siri, Inc. in 2010, launched Siri with next version of operation system of iPhone 4S in 2011 (Trappl, 2013). Unlike any other conversational agents or spoken dialogue system, Siri is deeply integrated into iPhone which enables the access to users' personal information like contacts and emails. Siri also has the integration of various domains and different web-services. These features make Siri a "virtual intelligent assistant". Siri is not just a chat-bot, it is designed to behave, or communicate like human (Luger et al., 2016). Powered with machine learning technology and natural language processing, Siri understands users' language accurately and responds quickly (Hoy, 2018). Things Siri can do include but not limit to making phone calls, setting alarms, controlling media playback, controlling smart home devices and making jokes.

Microsoft, Amazon and Google followed Apple shortly, launched their virtual assistants. Soon virtual assistants became the essential component in the phone (Corbett & Weber, 2016). Each virtual assistant has its unique features, but the core-features are in the same fashion (Hoy,

2018). What differentiates Alexa from Apple Siri, Microsoft Cortana and Google Now is that Alexa relies on voice communication using a screen-less speaker called Echo (Porcheron et al., 2018). Alexa only responds to the users with computer-generated voice (Whitenton & Budiu, 2018). As a virtual assistant built into a speaker, Alexa also offers different use scenarios from Siri and Cortana. Alexa aims at assisting users with their daily lives, with a focus on smoothing users' everyday routines by being a hub of home smart devices (Hoy, 2018). Alexa engages more in home-based activities. That is why Google Home and Amazon Alexa are called Home Virtual Assistants (Cho, 2018). Commands for Alexa include "Alexa, announce that dinner is ready", "Alexa, reorder paper towels", "Alexa, what's on my calendar today?" (Amazon, n.d.). Alexa also opened up the opportunities for third-party developers by allowing third-party manufacturers to make devices built with Alexa and third-party developers to exploit skills for Alexa (Janarthanam, 2017, p. 11). But the screen-less characteristic has raised some concerns about Alexa (Whitenton, 2016). Without a screen, Alexa only outputs a limited amount of information and possibly does a bad job at error prevention.

Virtual assistants are still in their early stages, many capabilities such as recognizing users' intentions and collaborating with other assistants are waiting to be fulfilled (Cohen et al., 2016), but it is undoubtedly the rise of virtual assistants that has brought the possibilities of performing tasks via voice controlling. Though the practicability of virtual assistants is still under discussion (Moore, 2016), virtual assistants have brought back the idea of natural language interaction and created space for future exploration (Luger & Sellen, 2016). Virtual assistants keep infiltrating into people's daily lives, yet academia has not shown enough attention on the use of virtual assistants in daily settings (Luger, & Sellen, 2016; Porcheron et al., 2018; Purington et al., 2017).

Next section introduces previous studies on virtual assistants and explains the establishment this study will build upon.

2.2 Previous Studies on Virtual Assistants

Since the emergence of Siri, researches on virtual assistants fit into three broad areas. These areas include evaluating the virtual assistants (Assefi et al., 2015; López, Quesada, & Guerrero, 2018), studies on using virtual assistants in specific scenarios (Carroll, Chiodo, Lin, Nidever, & Prathipati, 2017; Druga, Williams, Breazeal, Resnick, & Work, 2017) and studies on current challenges that virtual assistants are facing (Corbett & Weber, 2016; Cowan et al., 2017; Luger & Sellen, 2016; Moore, 2017; Porcheron et al., 2018;). The following sub-sections introduce each research areas on virtual assistants.

2.2.1 Evaluating and Comparing Virtual Assistants

Virtual assistants have made naturally communicating with computer realistic, which was always imagined in science fiction or movies. The big four virtual assistants include Apple Siri, Microsoft Cortana, Google Now and Amazon Echo. Each of them has unique features but share a lot of things in common. One of the virtual assistant research areas is to evaluate and compare each of these virtual assistants from user experience perspective, or the technological perspective.

Jadeja et al., (2017) claims that virtual assistants should be able to replace real human assistants in the future. They summarize characteristics of real human assistants, propose four perspectives to evaluate whether the design of virtual assistants is successful or not. These perspectives are user experience, information retrieval, linguistic and artificial intelligence.

According to Jadeja et al., user experience is the most critical one. Information retrieval perspective stands for the accuracy and speed that virtual assistants respond to the query of users. Linguistic is how natural the virtual assistants can communicate. Linguistic perspective has a matrix that contains 4 elements, quantity, quality, relation and manner.

López et al., (2018) evaluated the big four virtual assistants with two perspectives from Jadeja et al., (2017), user experience and linguistic. Participant was assigned with a script containing 4 kinds of tasks to conduct with virtual assistants and asked to complete a 5-point Likert scale for the naturalness and correctness of the virtual assistants. The results suggest that overall Google Now has advantages in naturalness and correctness of interaction compared to other assistants, while Amazon Alexa is more liked by participants in certain tasks like shopping.

Study of Cowan et al., (2017) takes user experience, linguistic, artificial intelligence as the evaluation perspectives. The researchers recruited infrequent virtual assistant participants who doesn't use virtual assistants regularly. The participants were asked to complete six tasks using Siri then interviewed about their views. The interviews reveal findings including inefficiency hand-free interaction with Siri, incapability of understanding the context and providing personal relevant service and issue of trust. As explained by the researchers, participants showed concerns about the privacy and data ownership. Participants do not have clues on how the companies will use the data. The issue of trust does not only exist in the data ownership, but also in the nature of collecting voice information. Alexa was discussed by Lei et al., (2017) about how Alexa may expose users in potential risk because Alexa can be activated by anyone who is calling it. Lei et al., argues that authentication on the presence of caller is necessary.

Through the comparison and evaluation, problems like privacy concerns and incorrectness of voice recognition were revealed. Comparing and evaluating virtual assistants do not take a big chunk of researches in this area. Next section introduces researches studying using virtual assistants in specific scenarios.

2.2.2 Virtual Assistants in specific using scenarios

The field of virtual assistant has put much attention on the use of virtual assistants in specific conditions or specific group of people. Researchers attempt to explore the potential of virtual assistant by extending the usage of virtual assistants to specific using scenarios. This section shows previous studies and explains the contribution they make to this field.

Assisting people with disabilities with virtual assistants is one of the directions. Pradhan, Mehta, & Findlater (2018) analyzed Amazon Alexa by collecting Alexa online reviews and looked for the keywords emerged from the reviews, then conducted interviews with visual impaired participants who use Amazon Alexa or Google Home. The findings were positive. All interviewed participants claimed that Alexa or Google Home is helpful. At current stage, the ecosystem of Home Virtual Assistants is not impeccable, the incapability to automate all home appliances causes frustrations to the users. While in future, it is predictable that virtual assistants will be more beneficial for people with disabilities with more devices connected. (Pradhan et al., 2018). Another group of researchers created a mobile prototype using voice interaction for motor-impaired participants to navigate through the interface (Corbett & Weber, 2016). Unlike the big four virtual assistants, the prototype Corbett & Weber created only allows pre-defined commands. The goal of the study was to improve learnability and discoverability by iterations of design. The researchers found that discovery-based learning could be promising for learners, but it still requires to be explored.

Similar studies were conducted on people with dementia and Alzheimer's disease (Rudzicz, Wang, Begum, & Mihailidis, 2015; Wolters, Kelly, & Kilgour, 2016). Rudzicz et al. used Wizard of Oz method in this study. They create a physical robot with a screen and a body. The movement of the robot was controlled by a remote operator, the capability to communicate was built in. The elderly participants with Alzheimer's disease were asked to complete a set of tasks following the robot. The results were disappointing, as the operator had to interrupt the interaction between elderly and robot to keep the task going. The issue exists in utterance of words of elderly people. A study conducted by Wolter et al. (2016) focused on people with dementia as well. Three focus groups were completed to reveal the participants' views on technology and virtual assistants that can be used for future design. The views vary on different participants, though most of the participants saw the benefits of virtual assistants clearly. The results also suggest the need of social bond for older people. Vardoulakis, Ring, Barry, Sidner, & Bickmore (2012) addressed this issue by a Wizard of Oz study. Remotely-controlled virtual assistants were placed at participants' home. The participants showed optimistic acceptance to the virtual assistants. Reis, Paulino, Paredes, & Barroso (2017) assessed the implementability of using existing virtual assistants to strengthen the social bond of elderlies. The researchers compared the big four virtual assistants on five criteria related to user identification and specific tasks. The accessed virtual assistants were able to complete most of the tasks, but not in notifying the social network updates. Overall Alexa is the closest one to meet all the criteria.

Another direction of researches is to apply virtual assistants to specific conditions such a collaboration environment. The results do not suggest a good fit of using virtual assistants in collaborative setting. McGregor & Tang (2017) created a low-fidelity prototype that listens to the communications in meetings and give support to the participants. The results were disappointing.

Participants rated the virtual assistants at relatively low satisfying level. The problem was on the capabilities to decide what's the right support to give. Porcheron et al. (2017) also suggest that interaction with virtual assistants in a collaborative setting is problematic.

Aforementioned studies have showed that virtual assistants can be beneficial in certain scenarios, such as assisting people with disabilities and strengthening the social bonds of elderlies, and less useful in collaborative settings. While these studies have not looked into using virtual assistants in a more common setting, the results do not necessarily apply to other situations as users' actions vary in different contexts (Corbett & Weber, 2016). Hence this study focuses on people without disabilities using virtual assistants in daily settings.

2.2.3 Current challenges

Though the concept of conversational agent has been existed since 1966 (Weizenbaum, 1966), and the emergence of Siri has shifted the ways people interacting with devices like smartphone and smart speakers, some fundamental problems are remaining undiscovered (Moore, 2017; Porcheron et al., 2018; Reeves, 2017). This question arises the concern of the term conversational interaction. Porcheron et al. (2018) conducted an study observing the use of Echo in home setting. They put Echo in 5 households for one month, observed the interaction between the families and Echo devices. Precheron et al. then argue that the term "conversational interfaces" is misnomered. They claim that the way people talking to virtual assistants is completely different to people talking to people. The constrains of conversational interaction is explicated by Moore (2017). The fundamental limits lie in the inequality between human and machine since the mismatch of capabilities of human and machine. The mismatch is because human and machine does not have shared experience and background. The mismatch can result in users not knowing what the machine is capable of and what to say to the machine.

Discussed by Reeves (2017), the fundamentals of virtual assistants are doubtful. Reeves explains the problem from four perspectives, action-formation, speaker selection, turn-taking and sequence-organizational problem. Action-formation stands for process of forming the actions. Reeves raises the questions “how do users know what they can do and how do users know what they should say?”. Study of Luger & Sellen (2016) uncovered how factors of virtual assistants influence everyday use. how users learn to use virtual assistants. They identified “humour of virtual assistants” and “understanding of context” as the factors influencing users’ use. The researchers also found that activities used by users to discover the capabilities of Siri and Google Now include playing and trial and error. But Luger & Sellen (2016) did not focus on HVAs, which the learning process can be different from Virtual Assistants as most of the HVAs are screen-less and used in different settings (Cowan et al., 2017).

HVAs required learning process because they do not provide true “conversational interfaces”, turn-taking is provided, instead (Reeves, 2017). Cowan et al. echo the statement and mention that users are uncertain about what the system could do. Users’ discovery process of HVAs’ capabilities remains unclear (Cowan et al., 2017). Cowan et al. used online surveys and semi-structured interviews to study infrequent people’s experience and use of Virtual Assistants. They discovered a few issues from infrequent users’ discovery process of Virtual Assistants. The issues include integration with apps, concerns about data privacy and social embarrassment.

Previous studies have set a tone for researching in the area of virtual assistants. Reeves (2017) brought up the questions about how users know what to say to the virtual assistants, however, Reeves provided a few considerations based on his own use of Amazon Echo, no empirical study has been conducted to answer his research questions, neither did Moore (2017). Luger & Sellen (2016) found the gap between user expectation and virtual assistants, but the

findings do not necessarily apply to HVAs because of HVAs' different use scenarios and screen-less characteristics. The lack of visual cues can influence how users learn to use HVAs (Pearl, 2017). Porcheron et al. (2018) attempt to add knowledge about interaction issues of HVAs in everyday use to the field. The focus of Porcheron et al. (2018) is to reveal how HVAs are embedded in social life. They did not look at the initial use of HVAs. According to Furqan et al. (2017), being able to discover the capabilities and limitations of a system is important for novice users. Few studies have looked at the discovery process of novice users. Building on the results of previous studies, this study aims at gaining a deep understanding of how novice users discover Amazon Echo Dot's capabilities in a home setting. Three perspectives will be used, the activities users engage in (Luger & Sellen, 2016), influencing factors of the design of HVAs (Cowan et al., 2017; Luger & Sellen, 2016) and influencing factors of the users (Cowan et al., 2017).

2.2.4 Summary

The history of virtual assistants and previous studies of virtual assistants are discussed in literature review. Next section introduces the methodology of this study.

CHAPTER 3. METHODOLOGY

This chapter includes the sampling method, data collection methods and analysis methods.

3.1 Process

To understand how the novice users discover Amazon Echo Dot's capabilities, this study used a combination of research methods to collect qualitative data. The participants were required to have no experience using HVAs. Before starting the study, the researcher had face-to-face meetings with every selected participant. During the meetings, the researcher introduced the study to the participants, explained what participants were supposed to do in the one-week-long study and distributed the Amazon Echo Dot to the participants. The researcher answered questions from the participants. The participants set up Amazon Echo Dot in the room where the participants wanted it to be. The researcher sent two surveys to the participants during the week. The surveys were inspired by diary study logs. After the first week, the researcher conducted follow-up interview with all the participants. During the follow-up interview, the researcher and the participant looked at the dialogue history of Amazon Echo Dot together. The dialogue history contains every single interaction between the user and the device. The surveys collected qualitative data from participants, they also were used in the interview results. The interview results were transcribed. Survey results and the transcripts were the collected data. The researcher used thematic analysis to analyze the data. The participants kept the Amazon Echo Dot for free as the compensation.

3.2 Sampling

This study used criterion-based sampling (Patton, 2015, p.577). The purpose of criterion-based sampling is to describe the characteristics of the group that meets the criteria. The study aims to provide an in-depth description of users' initial use of Amazon Echo Dot. Participants were required to have no experience with the Amazon Echo Dot or other HVA products. The selection information was included in recruiting information distributed via media platforms. The participants were from the greater Lafayette area as the researcher needed to talk to the participants face-to-face and handed Amazon Echo Dot to them. The sample size was limited to 10 because of the limitations on time and money. The participants were recruited by sending out study information via various media platforms (Wechat, email lists).

3.3 Data Collection

The data collection method of this study was inspired by the diary study. Diary study originates from the field of Psychology, then was used in multiple fields such as HCI and CSCW (Palen & Salzman, 2002). The current usage of diary study is to “capture activities that occur in real environments in relation to some kind of technology currently under investigation...” (Palen & Salzman, 2002, p.3). Early studies of diary study can be traced back to 1981 (Norman, 1983). Norman (1983) examined slips (a form of human errors) in different situations. The method has then been frequently used in HCI and CSCW.

3.3.1 Dialogue History and Survey

One characteristic of diary study is allowing participants to capture in situ data (Sohn, Li, Griswold, & Hollan, 2008). In situ logging is defined as when participants come across to a relevant activity, they log the activity (Flaherty, 2016). Sohn et al. (2008) conducted a study

exploring how mobile phone aids information searching by asking participants to log the information about mobile information needs whenever the participants needed information and had to look up via mobile phones.

This study did not use the data collection methods of diary study, but followed the spirits of diary study that capture the data in real environment. Amazon allows users to view user's every voice interaction with Echo Dot (Amazon, n.d.). The history provides a recording of the interaction and transcribed text if it is available. And the history also has time stamp on each interaction. Two distributed surveys served as supplements of dialogue history. The surveys had questions like "What activities have you done to discover the capabilities of Echo Dot?", which aimed at exploring user's activities other than interacting with Echo Dot directly, where the dialogue history could not capture. The surveys were also used as prompts for post-interviews. According to Carter & Mankoff (2005), it is suggested to use structured, question-and-answer based annotations to capture user activities.

3.3.2 Follow-up interview

After the researcher evaluated the data collected from the surveys, follow-up interviews was conducted with each participant. The follow-up interviews had predetermined questions. Interviews have been widely used in studies related to virtual assistants. Cowan et al., (2017) used interviews to explore the infrequent users' experience with virtual assistants. Rong, Fourney, Brewer, Morris, & Bennett (2017) used mixed methods including interviews to study the time expressions of virtual assistants. Pradhan et al. (2018) researched on how HVAs help people with disabilities via interviews.

Interview is an effective method to gather information (Berg & Lune, 2012, p.106). The researcher wrote down key words from each survey results, and used the key words as prompts

for the discussion in interview. Berg & Lune (2012) state in their book “Qualitative research methods for the social sciences” that “(interview is effective) particularly when investigators are interested in ... learning how participants come to attach certain meanings to phenomena or events” (p.115). Semi-structured interview is defined as a type of interview that including predetermined questions and special topics. An advantage of semi-structured is allowing the freedom to digress (Berg & Lune, 2012, p.112). As the interviewer, the researcher adjusted the wording and questions according to the context and extended beyond the predetermined questions.

The follow-up interview followed suggested sequence of semi-structured interviews (Berg & Lune, 2012). The audio was recorded, jotted notes was taken down. The interview began with demographic questions and asked participant’s general experience of Amazon Echo Dot, then moved to predetermined questions (Appendix B) about discovering activities and factors influencing the use, finally the researcher and the participant looked at dialogue history together, probing questions were asked. The participants were thanked.

3.4 Data Sources

The data sources of this study were survey results and transcripts from audio-recording in follow-up interviews, and jotted notes from follow-up interviews.

3.5 Analysis

This study used inductive thematic analysis to analyze the data. Thematic analysis is a method that identifies and analyzes themes (Braun & Clarke, 2006). The advantages offered by thematic analysis include the flexibility and detailed description of data (Braun & Clarke, 2012). Themes are the important information in relation to the research questions. Using inductive

thematic analysis means that the coding process does not require fitting the code to existing coding frame (Braun & Clarke, 2006).

The researcher followed six phases to conduct thematic analysis provided by Braun & Clarke (2016). The researcher transcribed the audio recording, combined the transcripts with survey results. Then the research started with familiarizing with the data, the researcher repeatedly read the data as a whole. After familiarizing with the data, the researcher generated initial codes across the entire data set. Some of the initial codes are shown in Figure 1. The researcher searched for relationships between the codes and collated them into potential themes. The researcher reviewed the themes, made sure the themes were related to the initial codes. Further more, the researcher named the themes and defined the themes.

The researcher rephrased the research questions based on the themes emerged from the data.

The original research questions were:

- I. What activities do users engage in to discover Amazon Echo Dot's capabilities?
- II. What factors of Amazon Echo Dot's design influence users' discovery of Amazon Echo Dot's capabilities?
- III. What difficulties do users encounter with Amazon Echo Dot at the initial use?

After reviewing the themes, the researcher found that the themes answered research question two, but from two aspects, the factors of Echo Dot, and the factors of the participants. Research question three is covered by research question two. So the researcher refined the researcher questions:

- I. What activities do users engage in to discover Amazon Echo Dot's capabilities?
- II. How does Amazon Echo Dot being screen-less influence users' discovery process?

III. What factors of the users influence their initial use of Echo Dot?

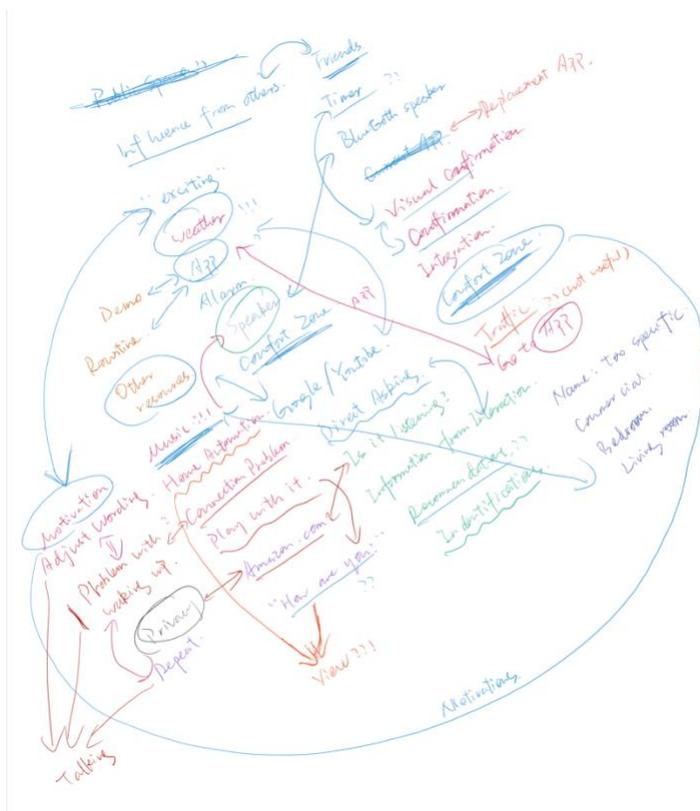


Figure 1 – Initial Codes

3.6 Trustworthiness

This study employed following strategies: generating alternative conclusions, making constant comparison (Patton, 2015), rich data and triangulation (Maxwell, 2013, p. 168).

Generating alternative conclusions stands for reviewing the data before settling on initial conclusions quickly (Patton, 2015, p.1371). The researcher revisited the data and attempted to explore other explanations for the initial conclusions. Making constant comparison shares the same core as generating alternative conclusions, while not only comparing the alternative explanations, but also the data don't fit into one category.

Rich data came from the surveys, dialogue history and follow-up interviews. Audio-recording of both methods will be transcribed verbatim to make sure the data are detailed enough for the conclusions (Maxwell, 2013, p. 168).

Triangulation requires information to be collected from diverse sources (Maxwell, 2013, p. 170). This study collected data from three methods, surveys, dialogue history and follow-up interviews. The researcher had access to the dialogue history during the follow-up interviews, which provided detailed data of the interaction between the participants and the devices.

3.7 Research Identity Memo

As Maxwell (2013, p. 166) states, the researcher's values and expectations may shift the conclusion of a study. Hence I, as the researcher, state my personal experience related to this study in this section.

I have always fancied the virtual assistants. I started using Siri since I got my first iPhone. Thought I was disappointed by the intelligence level of Siri, I still used it for simple tasks. I personally believed that Siri could be smarter in the future. Then Bixby of Samsung shocked me by its integrations with other applications in the phone and its voice recognition. I thought Bixby was the perfect virtual assistant, as I could complete many complex tasks with it, until I met Alexa. I got an Echo Dot in 2018. I had it set up as part of my home automation, I felt so relieved. Echo Dot captured every single command from me and responded quickly. I did not even have to say the commands loudly because of its sensitive microphones.

I found using Echo Dot or other virtual assistants easy and enjoyable, but my friends generally do not agree with me. I tried to explain the convenience of using a virtual assistant. I even let my roommate try using my Echo Dot for a week, then he gave it back to me at the third day, as he did not want to speak English, his second language, to the Dot. I started wondering

what stopped people from using virtual assistants. It is nature for me to interact with a virtual assistant, even the one without screen, as I know what it is capable of and the right way to talk to it.

As the researcher, I tried not to judge the conversations during the follow up interview, I tried not to influence codes and themes from the analysis with my own view. But my experience with virtual assistants might have still influenced the data collection process and results of this study.

3.8 Summary

The methodology chapter provides an overview of sampling method, data collection methods and analysis methods. This study used survey that inspired by diary studies and follow-up interviews to collect data, used thematic analysis to analyze the data. Themes generated from the thematic analysis are the outcome of this study.

CHAPTER 4. RESULTS

This chapter introduces the results of this study. The chapter begins with an overview of the results, then explains how the emerged themes are related to the research questions, finally shows the details of the themes.

4.1 Overview

4.1.1 Sample

The researcher spread the recruiting information via WeChat and asked the secretary of the department to send out my study information to the email list. A total of 10 participants were selected. Because of participants' schedules, not all interviews were scheduled at the 7th day of the participants' use of Echo Dot. The demographic information is shown in the table below. The information includes the participants' description of how familiar with technology they think they are.

Table 1. Participant Demographics

	Gender	Native Language	Education	Days using Echo Dot before Interview	Familiarity with technology
Participant 1	Female	Chinese	User Experience Design	7 days	Not confident in trying new technology
Participant 2	Male	Chinese	Mechanical Engineering	7 days	Pretty confident in technology
Participant 3	Female	Chinese	Biology	6 days	Would rate herself 5 out of 10
Participant 4	Female	Chinese	Finance	7 days	Pretty interested in trying new tech

Table 1 continued

Participant 5	Male	English	Civil Engineering	6 days	Would rate himself 6 out of 10
Participant 6	Female	English	Nursing	6 days	Pretty familiar with tech
Participant 7	Female	English	User Experience Design	6 days	Would rate herself 8 out of 10
Participant 8	Female	English	Bio Chemical	7 days	Pretty familiar with tech
Participant 9	Female	English	Hospitality Tourism Management	7 days	Very familiar
Participant 10	Female	English	Chemical Engineering	7 days	Would rate herself 7 or 8 out of 10

The interviews were between 37 minutes and 67 minutes. 19 survey results responses were collected. The transcripts have 115 pages in total.

4.1.2 How are the results related to research questions

The goal of this study is to find out how the novice users learn the skills of HVAs and what influences their discovery process. The research questions are:

- I. What activities do users engage in to discover Amazon Echo Dot's capabilities?
- II. How does Echo Dot being screen-less influence novice user's initial use?
- III. What factors of the users influence their initial use of Echo Dot?

The researcher kept the original research questions in mind throughout the study. The interview questions and survey questions were generated with the goal to answer the original research questions. During the interviews, the researcher asked the participants to describe how they learned the capabilities of Echo Dot and the researcher tried to dig as many details as possible. When looking at the dialogue history, the researcher asked probing questions on any activity the participant did not mention in the survey or in the beginning part of the interview.

Research question 1, 2 and 3 are answered by theme 1, 2 and 3 respectively. Following sections show the details of each theme.

4.2 Theme 1: The Activities Participant Engaged to Learn the Capabilities of Echo Dot

Theme 1 includes four sub-themes, learning through inquiring weather, learning via Alexa App, searching information online, asking Echo Dot and asking friend. These are the major activities the participants engaged to learn the capabilities of Echo Dot. Weather inquiring does not help the participant learn directly, but it led to other activities.

4.2.1 “Alexa, how’s the weather?”

Weather is the starting point for all the participants. The very first commands made by every participant were surprisingly consistent. Every participant asked about weather after they successfully set up the device. Weather inquiring also led to other activities because some participants encountered issues during the inquiring.

“That was my first question. I said checking the weather.” – Participant 1.

“I think the first thing I asked Alexa is what is the weather today in Lafayette, and then she answered.” – Participant 4.

The weather inquiring did not come randomly, but from the introduction of the tutorial of Echo Dot.

“...when you start off, start off the program, it shows some examples. I think the first is, what's the weather?” – Participant 5.

Asking weather was the entry point for every participant. Asking weather does not necessarily mean the participant wants to actually check the weather, it is also a way to make sure Echo Dot is working. This is explained by participant 10, *“I just wanted to know if I could, okay, that wasn't meaning to find out the weather. Later I did. This was just me trying to see if*

it's listening to me". Participant 8 did the same thing. "I think this is when I first turned it on and it was telling me like everything it could do basically...Um, so yeah, I dunno, I was trying to ask the weather to see if it's working".

Asking weather has also led the participant to explore more with Echo Dot accidentally due to different reasons. Participant 1 asked about the weather, found one word in the weather report she did not know, then she tried the translation function of Echo Dot to find the meaning of the word. *"...yeah, because I ask[ed] Alexa, what [is] the current weather? So Alexa said there is a flood warning on Monday and it says a word that I didn't know. I think it's like wind flowery or something. I want[ed] to know what the word is. [I asked her] what languages can you translate".* She also tried to change the unit of weather report. *"I asked about the current time, the current weather, I tried to change the, the like the unit of weather report from Fahrenheit to Centigrade degree."*

Other participants experienced the issue with the wrong weather location reported from Echo Dot, they tried to fix the issue with various methods. Participant 3 Googled the solution. *"That's another thing I had to Google actually cause whenever I'd say like 'what's the weather?' It would say in Chicago, Illinois, then list off the weather. So then a couple times I had to, like, ask about the weather in West Lafayette, Indiana and that worked. But then I found out that you could actually just set it to automatically saying what it was for this specific region. So that's when I went and set in the app".* Participant 8 experienced frustrations during the process of fixing the location issue. *"... it's like the weather, I had a little bit of struggle, which like you'll be able to see in that [dialogue history]like, um, the history with a knowing my location for some reason. Um, it automatically set my location to my, like, my permanent home address in Texas probably because that's my address on my Amazon account and I had to like figure out how to*

change that. Um, so yeah, I dunno, I was trying to get it to fix the weather, but it wasn't doing that. Um, so I guess it was just asking like what it could do cause I didn't know what to do. Um, so yeah, more, more attempts to, I said I'm in West Lafayette and she still didn't understand like how to change my location so I did have to do that manually. So yeah, it's just more of the same. I said Yikes.” – Participant 8.

Asking weather was not only how the participant started interacting with Echo Dot and exploring more with Echo Dot, it was also the most asked commands in the participant’s later use, partially because the weather in Indiana is unpredictable. The weather report of Echo Dot helps the participants to decide the outfits of the day.

“Okay. During the mid of the day, I go out, and want to check the weather again to decide my outfits, I ask the weather.” – Participant 2.

“I did ask about whether quite a lot, because Indiana weather is so unpredictable, by all means. Actually, I really liked the weather app because like with the weather feature she actually told you how it's going to be like it's going to stay this way. And she made it very simple because, some of the weather apps, it's very like it's hard to look through all the hours and like going up and down.” – Participant 7.

“I think the most useful is in the morning, in the morning when I'm trying to get ready and I'm super busy and I want to know what's the weather like to get dressed, um, time, temperature, um, and then checking the news too. I think that would help me in the future when I don't live in Indiana anymore.” – Participant 9.

4.2.2 Exploring with the Alexa App

Echo Dot has a companion app, Alexa. The app is a hub of related information of Echo Dot. It includes recommendations of skills, weather information, built in Apps. communication

center where the user can send message or call another Echo Dot device, and a media player. The participants found it helpful to learn the capabilities of Echo Dot by exploring the App.

Participant 1 learnt a few skills of Echo Dot through the App, while she also had difficulties understanding one of the skills, sending a message with Echo Dot.

“So I spent a few time checking the App, trying to learn what it can do. And I also, because it was a weekend, I did have some free time, so I play[ed] with the App for like a few times on that day.” “Let me see how did I know that. Uh, I see, oh, I think it's when I trying to figure out what Alexa can do in the app, it talks about how Alexa can talk to another Alexa. Because in the APP it talks about Alexa can contact or send a message to someone. So I'm, I wasn't sure what does send a message mean, is that trying to use my phone number to send an SMS message to like my friend's phone number from my contact list later? I think it probably means that it probably means send[ing] a message from one Alexa to another Alexa, which is associated with the account.” – Participant 1.

Participant 2 had very specific usage with Echo Dot. He considers Echo Dot as part of his home automation. He set Echo Dot to respond with weather, traffic information, news and turning on the lights to “good morning”. He found the App helpful for him to achieve his goals.

“The APP is pretty helpful. It will guide you through how to set up everything and even have a demo when you want to set a routine [to control the lights], like a series of movement. Like I said, in the morning you're going to see like, when I say one sentence, it's gonna respond with several other results. I set that up in the app.” – Participant 2.

Other participants tried what the App said.

“Um, as far as like just getting started with and learning to use it this past week, the app had like enough basic things for you to be able to learn” – Participant 3.

“Um, I think it just told me in like the APP, like, when I first downloaded the APP, it was like, ‘try a flash briefing’, and I was like, sure. Yeah, it was like suggested as one of the APP things. So the APP actually tells you a lot of like things you can try, which I think is nice. So yeah.” – Participant 8.

Participant 4 explored the built-in Apps provided by the Alexa App. He found a few built-in Apps useful, using the built-in App then became one of his frequent activities. *“Yeah. I found that [the built in apps], um, because we need to, use it through the App and I found that there are many different softwares on the App and that have it difference functions. It's pretty fun to explore it, those softwares” – participant 5. And participant 5 also learnt Echo Dot allows using multiple devices and link them together. “uh, you can access to your voice records in the app. and, uh, it seems as that you can associates more than one dot together and uh, you can label them with different room name. Like, one is in the bedroom and one is in the living room. So it, it would be very interesting to have many dots together”.*

When the researcher asked participant 10 how did she learn the skills of Echo Dot, she said that, *“mostly just through the Alexa App itself because it gives you a lot of suggestions right when the first open it up”.*

The Alexa APP is not only a place where the participants explore skills of Echo Dot, it is also a place where the participant look for solutions. Mentioned in last section, participant 3 found the issue with weather, then found solution buried very deep in the APP. Participant 2 wanted to change a setting of his routine and managed to change it through the APP. *“Oh yes, from the App. I was looking through the App, see anything I can do to fix it”.*

4.2.3 Google It

Searching online is the major method for all the participants to learn the skills of Echo Dot and find solutions to certain problems. Participant 1 went to Google to find out what Alexa can do, and found that Alexa can do translation in seven languages. *“I tried to use Alexa to do translation because I was searching for what Alexa can do in Google. Some people say, oh, that, like their daughter benefit[s], um, from like, Alexa translate[ing] the words for her and like, her daughter, her daughter can ask Alexa to spell the word for her, which is very nice I think is a good way to learn English or new English words”* – Participant 1.

As quoted in section 4.2.1, participant 3 used Google to find out how to change the default location of weather report. Participant 5 found articles through Google, *“I first read the description on Amazon. and I read some reviews on Amazon. Most of the people like it. And uh, I also read some article about, um, about how people, how to set up in different rooms”*.

Participant 10 said she always Google to find out the information she needs. *“I wanted to know how to make Pandora, like when I asked Alexa to play music, if I could make it so that she would go straight to Pandora instead of, I think Amazon music. Um, so I googled that and that was helpful.”* *“It was when I was originally setting it up. Alexa talked about a voice profile and I just skipped it or I listened to it and then it was like, oh I don't think I want to do that. So I didn't set it up then. And so I was trying to figure out how to set it up again after I had missed the message. And then I Googled and it said, do ‘Alexa, learn my voice’.”* – Participant 10.

Even when the participant has not solved the issue, the participant intend to search for solution via Google. *“Um, I'd probably Google my speaker's name and then like Echo Dot. And see if I can find like connectivity issues or something. Yeah. Um, or just like Alexa, Bluetooth connectivity issues and see if I can find someone else having the same problem online. Um, if*

not, I would probably go to Amazon, uh, like website and contact customer support, or like eco customers.” – Participant 8.

Participants also searched from YouTube, watched how other people use Echo Dot.

“I searched the web sites and I searched YouTube to see what do other people do with it.” “There are some videos compare Alex with Google home. Okay, basically, they are doing the same thing. So I don't know, uh, which, is the better? And also I, I, I see people like, uh, ordering, ordering on Amazon with this smart speaker and I don't think it's a very safe, so I would not do that.” – Participant 5.

“I did talk to some my friends and they showed me a YouTube video of this guy doing like crazy stuff with her. And like it's just a chain reaction. They're pretty funny cause it's, there's a whole thing on YouTube about like funny things they can do.” – Participant 7.

When the researcher asked participant 9 what she watched from YouTube, she said, *“like just YouTube video saying like introducing the echo dot. The difference between Echo Dot and the Alexa”*. Participant 9 also tried asking Alexa to tell a scary story after she watched YouTube videos. *“[I] watched YouTubers to do? Yeah. I think it was just YouTubers and I was testing it out”*.

Thought Google is a common strategy for all participants to learn and problem-solve, Google does not always have the answers. Participant 1 tried to find out how Chinese users use Echo Dot, as Echo Dot does not support Chinese.

“At beginning, Okay, Alexa can't speak Chinese, Um, then what I, what I can, so I'm curious like how Chinese user use Alexa because Alexa can't speak Chinese. As I said, for example, I cannot, I cannot ask Alexa to contact my friend because I saved their name in Chinese. So I'm just curious how people or how, like Chinese users who speak

Mandarin, how they interact with Alexa. Did they find like find a way around it or they just don't use those features.” – Participant 1.

She did not use useful information from Google thought. When the researcher asked her if she found anything, she blamed herself for not using the right keywords in search. *“No. No, I'm not sure, maybe I didn't type, like, um, uh, like good keywords so I didn't get some helpful information and I was all, maybe I was just too tired spending time doing research”.*

4.2.4 Ask Alexa or Ask a Friend?

Asking Echo Dot is a direct way to explore its capabilities. All participants have tried talking to Echo Dot directly to ask about the skills. Though most participants did not find talking to Echo Dot very useful, as the information provided by Echo Dot was already listed in the Alexa App or the tutorial booklet, or Echo Dot does not have the answers.

“Um, it [talking to Echo Dot] kind of went through like more of the basic stuff that was already listed in the booklet and then it had like a couple of extra things and then it said that I could like it told me to... At the end of the list that I went through, it was just kind of, it said to go ahead and look at the APP for more options of what to do.” – Participant 3.

“I actually just learned there is a, like I, I usually say, ‘Hey Alexa, what can you do?’ Or like I said, hey, ‘what features, what new features do you have?’ And she's like, ‘I don't understand you’. I'm like, of course you don't.” “I actually talked with her about where it's like the tutorial, I'm learning to go through it, but it got even more, like, I started going through like the daily, like music and stuff like that. And once I got bored, I would ask her like, what can you do you? She's like, I can help... Pretty much every time I asked her, she always brought up cooking somehow.” – Participant 7.

“Yeah. So I started using the lists cause I asked ‘where can I take notes’ and she [Alexa] didn't like have a response and then I Google, does she have like a way to just take notes?” – Participant 10.

Participant 8 did learn something from asking Echo Dot. *“I was just, like, trying to figure out more random, anything she could do, just kind of talking to her randomly. Um, and then she said like, you can ask me about a quiz or something. I'm pretty sure at one point she was like, you can ask me for the quiz.”.*

Comparing to learning from asking Echo Dot, learning from the friend is a more helpful way for the participants.

“Yeah, because I kind of know it, what it does before I got this dot, because a lot of my friends use Alexa and Google Home so they are pretty similar to each other and I basically already knew what it does.” “Well, at first I learned what Alexa does from my boyfriend because he got a Christmas gift from his brother, like, which is an Alexa Echo Dot. It's pretty cool. And when he first set up in his bathroom and tells that to do stuff, I learned a lot of function from him. And also I have other friends who use Alexa as well. So they told me a lot of stuff and then I'm probably going to look up on Google or like just go through Alexa instructions on the Amazon so I can find out more functions.” –

Participant 4.

“I had a friend tell me that I could do that because I guess her parents have it.” “Um, I, I've learned a lot from my friend who also has one. She's the one that told me about the games like jeopardy and that I could connect it to my Spotify. So I think that she told me a lot of things I could do, but some things it's like asking and seeing what all it could do, what it couldn't do. That's how I learned a lot.” – Participant 6

“I did talk to some my friends and they showed me a YouTube video of this guy doing like crazy stuff with her.” – Participant 7.

“My friends said that it's really good to use if you like want to power your house. They have to use, they asked me to buy a, um, uh, command, um, a dock or something that connects the house lights to Alexa or to Echo Dot. And then this way you can command it to turn off all your lights.” – Participant 9.

Even for those who have not yet learnt anything from their friends, they would seek help from friends in the future.

“Yeah. But I would rather learn from my friend, like an actual person, coming to my house help me with learning all the new features because that is easier.” – Participant 1.

“I haven't, I know that like in the future if I wanted to know more about it, I have a friend who has one that I could ask.” – Participant 3.

The reasons were explained by participant 1. *“My most preferred way to learn the those is having my friend come to my home, tell me about Alexa. I think the reason is I don't want to spend time and do the research online. I'm not sure whether I can find, like, good articles or blogs about this. I don't want to spend time reading those text, heavy text and I don't want to try to ask Alexa what she can do because Alexa can say a lot of things at a time. I couldn't learn them. Also Alexa is not a real human beings so sometimes she doesn't understand what I need so I may feel a little bit, like, upset or disappointed. But if my, if my friend is coming to my house helping me, helping me learn things, because he or she is a real human being, you can always ask questions and they all understand what you need because they are a, they are just as another user.” - Participant 1*

4.3 Theme 2: The Influence of Echo Dot Being Screen-less

Theme 2 contains the influence of Echo Dot being screen-less. The fact that Echo Dot is screen-less makes some activities more difficult. Talking to Echo Dot and lack of confirmation are the main issues. Privacy concern is also discussed by the participants.

4.3.1 Talking to Echo Dot is not Easy

Most participants experienced different issues talking to Echo Dot. One of the issues was the incapability of Echo Dot to understand participant's commands. The participants needed to repeat themselves or change the wording.

“So when the alarm clock actually alarm, so, I say ‘Alexa, I got it’, I thought it would understand so I can turn it off. But he did not understand that as turn[ing] off the alarm. So I need[ed] to say ‘stop’ to actually turn off.” – Participant 2.

“[The experience was] usually pretty well, although the voice recognition was a bit flawed, I guess. When dictating texts, it wouldn't have the right words on there sometimes. Which is understandable.” “It was, I think, the voice recognition on the dot. Just like a bit [flawed], um, I couldn't figure out how to tell it to delete the reminder.”

“Um, probably like just recognizing like what you're saying whenever you're speaking to it, because there were a couple of times that it had trouble understanding me when that unknown areas. Being able to differentiate, like what you're saying from like one word to another, understanding that would be a good improvement.” – Participant 3.

Talking to virtual assistant is never easy. But being screen-less makes it even more difficult because Echo Dot only accepts commands via talking. When the participants want to achieve certain goals, they need to repeat or rephrase themselves to make Echo Dot understood.

“Sometimes I have to repeat myself, but it's nice when it actually understands what you're saying. Uh, sometimes they'd have to speak slower and louder for it to understand.” – Participant 6.

“I had a few issues of, like, it didn't give me enough time cause I was trying to, like, think about what I would call it. Cause it's one of those things of like, how would I phrase it so it would understand.” – Participant 7.

“I was excited at first and then started getting annoying trying to get the attention because I would say something and it doesn't register and I have to repeat it over and over again.” – Participant 9.

Participants also had issues with waking up the Echo Dot. Participant 5 found it difficult for him to pronounce the wake up word, Alexa. *“Sometimes it's hard to wake it up and uh, and you need to, do you need to say ‘Alexa’ to wake it up. I would say ‘Alexa’ is difficult for me to pronounce it right. So I failed a couple of times.”*

Participant 7 was annoyed by using the wake up word every time talking to Echo Dot. And Echo Dot happened to wake up without anyone talking to it. *“Um, I was kinda disappointed how many times I've had to say Alexa gets so frustrating.”* *“I think that's because she recognizes my voice. So that was very, um, funny. I've had her turned on when I'm talking about different things also and not like the command ‘Alexa’ but I think Google does that too. Yeah. And just like I didn't, I wasn't talking to him. thanks for butting into the conversation.”*

Requiring too specific commands also frustrated the participants.

“And that's another thing that we had a problem with was when we tried to play a playlist on Spotify, we had to be like super specific with it.” – participant 6.

“Um, and I would need to know the exact name of the song in order to play on Amazon prime music.” – Participant 9.

“I mean, I think that the only real issue I had was accessing the skills just because it required such specific wording.” – participant 10.

Participant 10 is the only person who had smooth experience talking to Echo Dot. *“I mean, I like now that I've gotten familiar with like sort of the way that it likes to hear commands phrase that it's very receptive and immediately you will just do whatever. I also liked. At some point it asks me ‘do you want to like, do you want me to give you shorter responses?’ Um, and so it was nice that it was like, I dunno, offering to improve the experience like as I was using it.”*

4.3.2 Where is the confirmation

Participants wanted confirmation of their activities, but Echo Dot does not support it very well. There were two kinds of confirmation the participants searched for, visual confirmation and voice confirmation. In some situations, the participants wanted visual confirmation of what they have done. Examples are:

“In my shopping list, I, I think I'm visual learner, so that's, I'm telling Alexa to write it down, but because I didn't see it, it was like written in my memo, I think it doesn't help me remember. I already, I already put this in my shopping list. Like maybe later I was thinking, Did I put that in my shopping list? I think for me it's still better to just either write it down or typing it on my phone by my hand.” “It tells me like the alarm clock it was set up or something, but um, but I think it's already, it has already been a long time, but Alexa didn't say anything so I saw[looked at] it.[I thought] it's already, it's already 16 minutes or more than 16 minutes. Because I was baking stuff, the time can't pass that

time. So I wanted to just check with Alexa, to make sure. Just like if I set the alarm on my phone, I will look at the screen a few times before the alarm.” – Participant 1.

“Trying to set my calendar up with it. Um, put a reminder just cause I can't see what I'm doing, I'm only saying it into their machine and I can't see what I'm registering into the APP. That kind of gets me annoyed. Um, I think I like to have, like if it's on my phone, the APP, I can see what I'm registering into it or I can, like, do it myself and put it in. And I have to say it [to Echo Dot].” – Participant 9.

There were also situations which were not caused by Echo Dot being screen-less, but not providing timely feedback to the participants.

“So I'm like on like whenever you set one, it says that it might take a few minutes for it [routine] to register to the dot. Okay. So, um, like after I said it, I tried it a couple times to see if it was a routine or not. And it took about two minutes, I think. First registered then repeat it.” – Participant 2

“I mean with setting reminders, it was difficult to do it. Um, like by activating the dot, like saying Alexa and then setting it. So I had to go into the APP to, to make sure that it had the right reminder.” – Participant 3.

“But the idea of like some of the functions like that, I would use it. She's like, I can set an alarm. So I really wanted to do with the alarm function. But the thing was, I didn't know how loud her alarm was. I didn't know what the song she was going to play and she directed me back to like the Alexa App or you know, the APP. And I was just like, okay, I could do that or I could just keep the alarm on my phone.” “Have to ask how much time and I canceled the timer because I was just like, ok I'm not sure how the, how the thing

would be like do I do, do I risk it or do I just set something on my phone? It's silent but I can actually see.” – Participant 7.

4.3.3 Privacy and Security

Three participants brought up their concerns about the privacy and security. Participant 5 unplugged his device because he thought Echo Dot is always listening. *“Because is it is always listening and so sometimes I just cut the power.”* Participant 9 talked about what she heard from her friends. *“It's scary. So my friend during, yeah, her conversation in the room, like she was just talking to her boyfriend about the Alexa and how they gather information and a send to the headquarters and listen to you at a time. And Alexa just responds with ‘I don't, um, I don't send information to headquarters.’ It was scary. So she unplugged it.”* And participant 9 also described Echo Dot as not loyal, as she does not want Echo Dot respond to other people. *“[Echo Dot is] not loyal, because she responds to another person. I don't know how she, does not register my voice.”*

Participant 7 talked about how she was worried about purchasing with Echo Dot. *“And that's another thing as I felt a little weird about ordering something on the dot. So I didn't, because I was just like, my credit card, you know, what if like, like my roommates, they ordered something online. I know you can put like a password or something on it, and it was one of those things of like, I don't know how secure this is. And nowadays everything's listening.”*

4.4 Factors of Participants that Influenced the Initial Use

4.4.1 Language Barrier

Three participants had issues communicating with Echo Dot due to their native language being Chinese, which is not supported by Echo Dot. Participant 1 could not call or text her

friends because her friends' names were all saved in Chinese. She also abandoned using the shopping list supported by Echo Dot, as some names of the vegetables and foods are in Chinese.

“So which [Echo Dot does not support Chinese] is not a good thing. Like I mean, uh, like if I say my friend's name in my contact list, most of the time it will be their Chinese names. But if Alexa can't speak Chinese, I can't speak Chinese to Alexa as well. So I can't like ask Alexa to, like, contact someone.” “Another reason is, maybe sometimes I just want to buy the product that doesn't have an English name or I doesn't know the English name, like Chinese vegetables. How can I ask Alexa to save [to shopping list]?” –

Participant 1.

Participant 5 found it difficult to pronounce the wake up word, “Alexa”, correctly.

“Sometimes it's hard to wake it up and uh, and you need to, do you need to say ‘Alexa’ to wake it up, I would say ‘Alexa’ is difficult for me to pronounce it right. So I failed a couple of times.”

Participant 3 had issues with pronunciations as well. She tried to text with Echo Dot, Echo Dot misheard her words. “Yeah, that [mishearing the words] happens. But I don't think that's her fault, I think it's my fault because I didn't pronounce it perfectly well.”

4.4.2 Participant's Familiarity of Technology and Their Motivations

During the first week of participants' use of Echo Dot, participants' motivation influenced how they use and learn to use Echo Dot. The participants' motivation levels were also bonded to their familiarity level of technology. Some participants do not want to try the new skills of Echo Dot, either because they were not highly motivated, or they did not think they needed those skills.

“Because I'm not really a people who are interested in trying new technology. Yeah. So for me, if I'm crazy about new technology, I love to try it. I would definitely try those out

as well. But, I am more comfortable using the current or most popular technology. So I'm not highly motivated, but if I'm really free, I mean, if I'm a student I do have lots of free time, I will try to rest [of the Echo Dot skills] out.” - Participant 1.

“Uh, I was pretty motivated at the beginning, but you know, the, the descend gradient is pretty large. So, yeah, I'm not a very motivated now.” – Participant 5.

“Um, I dunno, probably like pretty motivated in the first couple of days. Um, and then, like, as I, like, kind of stopped using it, not as motivated as it had, as I had like kind of a connection problem. But like I said, once I troubleshoot them I think it'll be fine. I just haven't really had time to do that. So it's, it's kind of a time issue, kind of app issue, but it is what it is.” – Participant 8.

“I'm motivated to use it. Um, I mean for the, the few things that I like I'm still doing today. Um, those things, like super easy and I'm very motivated. I use them now just regularly. Um, the first day I was really interested in finding new skills and new things. Um, and so I tried a few things the first day, but just kind of found that they were, I guess a little bit more trouble to use. Um, but I was, um, I mean I'd still definitely be interested in finding out more. Um, but yeah, as far as the things that just sort of pop it up naturally.” – participant 10

Some participants know there are more skills of Echo Dot, but they do not see the reasons exploring the skills. When the researcher asked participant 3 what skills does she want, she answered, *“I'm not sure, cause I feel like something that, I'd say, is probably already does, that I don't know what it does. Um, yeah, I can't really think of anything that I really want.”*

Participant 7 summarized the reason she does not want to explore more skills.

“So she can do a lot. It's just my comfort zone because it's like if I asked her to like call, oh, I have asked her to call people, too. I don't usually use speech to text because I find it what you say usually gets translated, and you're just like, okay, now I have to retype that anyway.” “I didn't go in detail about that [skill recommendations of Echo Dot] cause I was just like, again, my comfort zone was here. It's interesting because I don't think I really explored more. I just explored a little bit outside my comfort zone but that's it. At some point I was just like my phone can do that.”

Two participants were very motivated during the first week of their use. Participant 3 found it interesting to interact with Echo Dot using voice. *“It's definitely interesting using it and it's, um, it's kind of fun, honestly. Yeah. It's fun for me to use like the reminders and everything is just kind of nice to be able to, say that and then have it be, um, I guess, vocalize, about it.”* – participant 3.

Participant 9 was very motivated during the study. *“I was really motivated. I was really excited to use it because my friends always used it. And I really liked having that easy access of asking things. I also used it to ask to tell me about stories, like scary stories when I was bored. Um, I watched it in the video and I was intrigued by it, so I used it. Um, yeah, I think it's pretty cute, capable of lot.”* – Participant 9.

4.4.3 Previous Experience with Virtual Assistant

The participants were required to have no experience with HVAs. But all the participants have used traditional virtual assistant such as Siri and Google Now. Some of the issues or the usage of traditional virtual assistant were transferred to Echo Dot. Participant 1 had issues learning using Siri due to the learning curve. She faced the same issue during the use of Echo

Dot. *“Um, because I'm not sure how to achieve my goals. Uh, there is large learning curve, learning curve. I just don't want to spend timing figuring out how to do those.”*

Most participants also followed the pattern of how they used traditional virtual assistant. Participant 4 used Siri mostly for calling or texting, she did the same thing with Echo Dot. *“most of the time it's just call my friends or text people like doing homework or doing other stuff when I'm not on my phone. So I can just say something to Siri so that she can stuff, she can do stuff for me”*. Participant 5 also did what he asked Siri, with Echo Dot. *“I asked a couple questions like some questions about its identification. That's, ‘who are you’. I and also some questions , some questions about the history. And also I asked the about the whether, the temperature. I have done the same thing on the echo.”*

Participant only used virtual assistant for very specific tasks, “Back in China. I have the virtual assistant on my car. And, uh, that's only time I use it [for navigation]”. He used Echo Dot for his home automation. All his activities were concentrated on this.

4.4.4 Participants' views of Echo Dot

The participants had different opinions about what Echo Dot is after the first week of use. Their opinions also influenced how they used the devices. Some participants consider Echo Dot as a personal assistant. They think Echo Dot is capable of getting some simple tasks done, not the tasks a real human can do.

“I say it's, for me it's more like an assistant because she can remind me what the current time is, what kind weather is, she can play some background music and I uh, and she can set an alarm for me. But I don't expect her to help me get things done because when I try to get things done, I didn't rely on her. I didn't use Alexa to work on like the, to do list or

shopping list and I didn't use Alexa to set reminders. So for me it's just like assistant.” – Participant 1.

Participant 1 also talked about how Echo Dot looks like to her Bluetooth speaker. “Yes. *Uh, another thing, I'm not sure, Bluetooth player, yeah. I'm not sure whether Alexa can work as, uh, Bluetooth player or not. So in my mind, I think it can. It just looks so similar to my Bluetooth speakers. And that's why sometimes I think, why Alexa can't play music from my phone, or maybe I just didn't figure out how to do that. I wish it can do that. Then I can play music on my phone and I can get rid of my Bluetooth player.”* Participant 5 had similar thoughts, “*it's not that intelligent. It's just a speaker.*” So did participant 6. Participant 6 considers Echo Dot as a speaker that can also function as a convenient tool. “*I used it a lot as a speaker for my music, but I think that's, it's a good tool for when you have to ask things”.*

Participants generally did not have very high expectations of Echo Dot

“So I wouldn't expect it to, to do something it can't, for example, like something like an actual, like, a robot or a person can do. So I know it's only a virtual assistance so it can only do whatever you commanded to do, like on a virtual basis.” – Participant 4

“I think it's like all the other artificial intelligence. It has a name, I mean it's not a human being, it doesn't have, like, thoughts. But it speaks to you every day. So I kind of treated it like, I mean, like virtual assistant. It's not a real, it's not a person, but it still does a lot of stuff. More like a personal assistant, personal assistant. Yeah.” – Participant 4.

“It's, it's smart, but it's not smart. Smart to a point, let's just put it that way. And then the human has to be smart because they have to know how to phrase things and talk to it actually. And then you look like a crazy person talking to yourself,” – Participant 7.

Some participants also saw the human aspect of Echo Dot through the interaction. *“And then if I was leave for class, I would say, ‘okay goodbye Alexa’. And she would come back and say goodbye or some... That made me so happy when I heard it. Cause it's like the weird, you're away from home. Like you don't have anyone say goodbye to or whatever. She said goodbye back. And I was like that. Or like I said something this morning of like good morning or something and she responded. It was kind of funny but it was a small thing, interactions like that that made me smile”*. Participant 3 liked the Echo Dot said “You’re welcome”. *“I mean it's kind of like a nice touch for it to respond to that with the ‘you're welcome’ or anything like that, like a person would.”* - Participant 3.

Two other participants view Echo Dot as part of the home automation.

“I think it's part of my home automation, that's how I treat it. The reason I want her in the first place was because I need to control the light switch and the light is kind of awkward position. I rather have Alexa to control by voice or I have to get remote control.” – Participant 2.

“Um, I dunno. I think it's very useful for connecting to other devices. And controlling other devices. And I think it has a lot of capabilities when it comes to like the whole like smart home concept. I really think that that's where like most of its power lies.” –

Participant 8.

4.4.5 The Decision on the Placement of Echo Dot

Most participants set up the Echo Dot in their bedrooms. Some participants live in dorms, they had no choice where to set up the device. Those who live in a house chose the placement based on where they present the most.

When the researcher asked participant 1 why she set up the Echo Dot in her bedroom, she answered, *“Because I think the most frequent case will be I ask Alexa what time it is while I’m lying on my bed. So it has to be in my bedroom. I don’t do lots of the cooking or something in the kitchen, but I will go to bed, everyday. I will wake up everyday. I need to ask Alexa to check the time. I think that is the most frequent use case. So I, I decided to put it in the bedroom.”*

Participant 1 also talked about one time when she was cooking, she went to the living room, which it is closer to the Echo Dot, to set up a timer. *“I think I went back to the bedroom. So that I think Alexa can definitely hear me. If I’m in kitchen, I’m not sure because I think maybe Alexa can hear me, but I cannot hear Alexa. So I decided to go back to the living room, which is closer [to the Echo Dot] so I can hear what Alexa is saying.”*

Participant 4, 5 and 7 also had their Echo Dots in their bedrooms.

“Because my bedroom is where I study and then I go to sleep and where I wake up. So I think it’s easier for me to like set an alarm or like there was, or set on a timer or like put stuff in my shopping list or like to do lists.” – Participant 4.

“Uh, yeah. Um, I like to read on my bed and this is one reason. Then another reason I can use it as a clock, an alarm clock to wake me up.” – Participant 5.

“Cause I think the dots more for like relaxing and enjoying, and I do that most in my bedroom. So that make senses of putting in my bedroom. But like some of the things she used for, I could definitely see it being used in the kitchen. Then you have the dilemma of do you put in the kitchen or you put in your bedroom because you can’t move her around that much.” – participant 7.

Participant 7 did not put the Echo Dot in the kitchen, but Echo Dot suggest the ability to recommend recipes for her, which annoyed her. *“And once I got bored, I would ask her like,*

‘what can you do you?’ She’s like, I can help...Pretty much every time I asked her, she always brought up cooking somehow. Yeah. She’s like, ‘I can recommend recipes for you.’ I was like, is it known for cooking? Because like in my, in the app, you specify where it is in your house. She knows it’s in the bedroom. Why you bring up cooking to the bedroom? That’s why I was confused about.”

Participant 10 set the Echo Dot up in the kitchen at first, then moved it to her bedroom. *“Um, well, I set it up in the kitchen, but I moved it to my bedroom because I found that I spend most of my time in there rather than in the kitchen.”* – Participant 10.

Participant 2 used Echo Dot as the center of his home automation. *“Um, it [Echo Dot] is on the bar between the kitchen and the living room. It’s kind of a center location of my apartment. And it is pretty near the power.”* – Participant 2.

Participant 8 set her Echo Dot in her living room. *“Like that’s where I spent most of my time. Um, I considered setting up in my bedroom but um, I like, I spend most of my time and my living room cause that’s where like my couch it and my like tables are. So I set up on my table and then I can still hear it and usually talk to it from my bedroom as well. If I’m, if I’m not too quiet, I can usually talk to it from the bedroom. So that was a pretty good place for it. And then in the living room too, I can talk it from the kitchen. So that was nice so that I could use it while I was cooking.”* - Participant 8

CHAPTER 5. DISCUSSION AND CONCLUSION

This chapter includes the discussions based on the results, along with the insights for the HVA design provided by the researcher.

5.1 Discussion

This section provides the researcher's interpretations of the study results. The results are discussed from four perspectives. The first perspective discussed the proper help that Echo Dot and Alexa App should provide to the users. The second perspective describes the feedback Echo Dot is supposed to give. The third perspective explains the importance of building proper user expectations. The last perspective is the privacy concern.

5.1.1 The Proper Help that Echo Dot and Alexa App Should Provide

As described in the last chapter, participants in this study prefer referring other resources to learn and solve problems, such as searching online, asking friends, than asking Echo Dot. This is due to the limited help that Echo Dot can offer.

The nature of voice-interaction only has determined the lack of pedagogic commands, which teaches the user about the functionalities of a system (Cohen et al., 2016), on Echo Dot. Most used commands on Echo Dot are invisible commands, that no direct indication is available for these commands. The invisibility results in the Gulf of Execution (Norman, 2013), that the users do not know what can be done. The gulf will not exist if the Echo Dot can truly simulate natural voice interaction and be intelligent enough to answer every single questions from the user. But the reality is that Echo Dot is very task-oriented (Luger & Sellen, 2016), novice users still need to learn the capabilities of Echo Dot.

Amazon puts limited efforts making the invisible commands “visible” to the users, instead, Amazon hopes to teach the users basic information through the tutorials and booklets and let the users discover more skills via the accompany App. But the information included in the tutorials and booklets is rather simple. For those users who want to access more skill information via asking Echo Dot directly, they may be disappointed because Echo Dot would answer with no additional information other than the basic skills that can be found from the tutorials and booklets. As a voice-enabled device, Echo Dot is not allowing users to learn through the voice interaction, but hoping the users learning with the App and other resources. The lack of transparency can cause difficulties in user’s learning and problem-solving (Luger & Sellen, 2016). Amershi et al. (2019) put “make clear what the system can do” as the very first design guideline of human-AI interaction. One way to achieve the clarity is to provide progressive assistance for voice interface throughout the interaction (Yankelovich, Levow, & Marx, 1995) rather than giving almost the same feedback every time, which is what Echo Dot does right now. Progressive assistance means giving short answers at the beginning, offering more help if the questions persist. If a user asks Echo Dot “what can you do” a couple of times, it is obvious that the user needs more information than the simple answers.

Echo Dot should also provide contextual help (Amershi et al., 2019; Cooper, Reimann, Cronin, & Noessel, 2014, p.385) where the help is needed. Corbett & Weber (2016) found it helpful implementing contextual help in a virtual assistant interface. According the guidelines for human-AI interaction (Amershi et al., 2019, p. xx), the virtual assistant needs to allow users to find information related to current tasks and environment. As mentioned in last chapter, several participants had issues asking weather for the first time. The participants attempted to search for

contextual help but it is not supported by Echo Dot. The issues break the use flow of the participants, and require the participants to jump to the Alexa App.

Alexa App has a page of overview information and random recommended skills for the users, as shown in figure 2 (left). The App also contains more detailed information about the

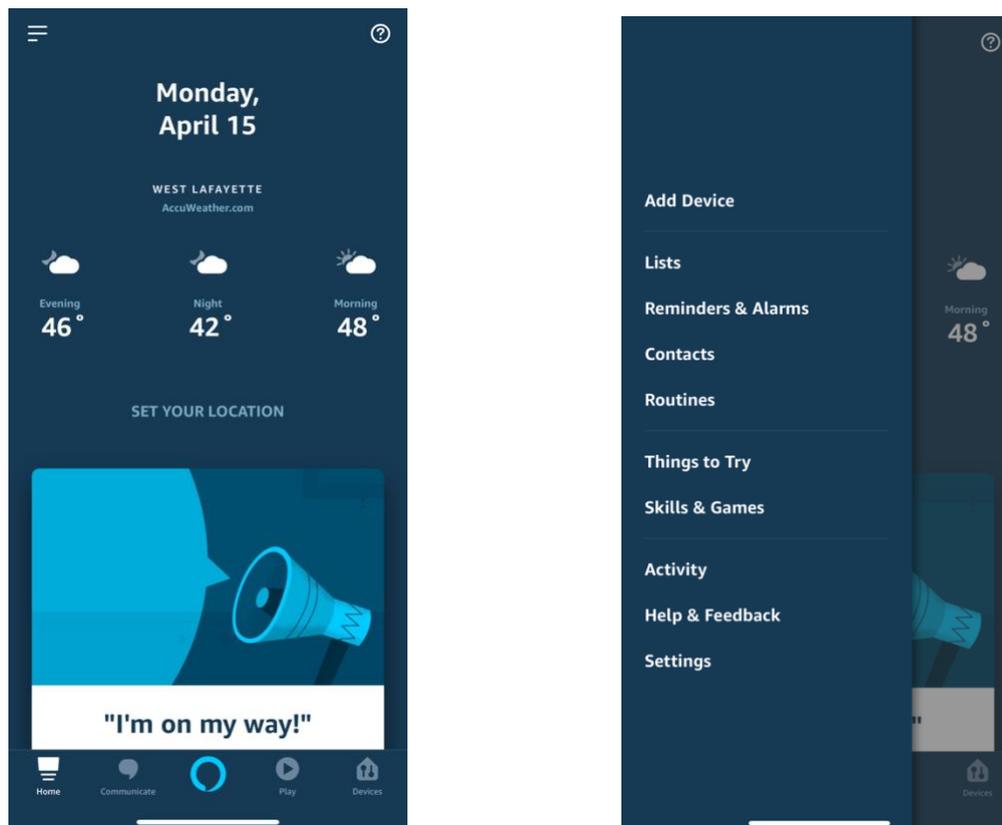


Figure 2 – Alexa App

skills, but the detailed information is buried under the sub-menu of drawer, “things to try”, shown in figure 2 (right). Discussed by Cooper et al. (2014, p.385), the use drawer hides the functionalities, causes problems when the user searches out for certain things. An ideal way to improve the Alexa App is to have an individual tab for the skills. It is also necessary to add search function to the app. Search is the most important mobile activity in a mobile application (Cooper et. al, 2014, p.544). The participants in this study had difficulty finding out a specific

skill, which resulted in giving up looking through the skills of Echo Dot. Amazon should also consider adding different media forms. An article introducing the skills or a video explaining more detailed information can be helpful for those users who want to learn.

The contextual help should not only be provided during the interaction between the user and Echo Dot, but also in the Alexa App, where the skill recommendations show up. In the Alexa App, the skill recommendations are listed in the home page (Figure 1, left). These recommendations are not related to user's activities or behaviors, they are based on the popularity of the skills. In this study, the participant had the case that Echo Dot suggested unrelated skills even knowing the placement of the device, which frustrated the participant. Users have different choices for the physical placement of devices, the placement influences how the users use Echo Dot (Sciuto, Saini, Forlizzi, & Hong, 2018). The participants in this study chose the placement because of particular goals or where the participants present the most. Echo Dot and Alexa App should personalize the recommendations at least based on the placement. Cohen et al. (2016) indicate that personalization should involve many factors of users, including user's identifications, goals, intentions and so on. The personalization can be achieved by using machine learning. The system should be learning through observation, being told and user demonstration, which Echo Dot already has access to. It would be ideal for Echo Dot to use the information that it can access and convert the information to the actions that can ease user's activities. This brings the concern of privacy, which is discussed in section 5.1.4.

Besides the Echo Dot device and Alexa App, Amazon should also prepare a right place for the users who search for definitive answers. It is common that the users turn to the help of traditional online search when they have questions or expand their knowledge of a product

(Cooper, et al., 2014, p. 393). A place that includes common problems and corresponding solutions can be very helpful.

5.1.2 Timely and Accurate Feedback for the Users

Failed feedback has caused frustrations to the users. The participants in this study kept searching for visual or auditory confirmation when they did not get proper feedback. According to Nielsen (1994), the usability of a system is largely dependent on the visibility of the system status. The visibility requires timely and accurate feedback and feedback that keeps the users informed what happens. However, Echo Dot being screen-less has minimized the possibility to provide visual feedback, the way Amazon deals with it is to use LED lights on the shell of the Echo Dot device. The LED lights show the status, interaction and notifications using different colors and animations (Martin, 2018).

Obviously the feedback provided by Echo Dot does not fulfill the requirement to show the status of the system. The participants in this study have troubles even in the most basic activity, realizing Echo Dot is listening. Summarized by Porcheron et al. (2017), among visual, touch and voice feedback, users still rely upon visual feedback more. This explains why the participants in this study still went to the Alexa App to check the results of their activities after they completed the tasks with Echo Dot. Adding a screen can always be a good solution, but most use cases in this study happened when the participants were hands free, occupied by the main tasks, and the participants want to use Echo Dot as the aid. Requiring visual attention could distract the participants from the main tasks. Thus it is significant for the industry to explore the possibilities to create a new modality of feedback that relies mainly on voice, assisted with LED lights.

The researcher suggests Amazon to enable users to adjust the frequency of auditory feedback. As now Echo Dot allows users to shorten the responses from the device, it is also

reasonable to allow users to get more detailed responses. Taking reminder as an example, the participants in this study had problems knowing whether the reminder had been properly set up or not. It could be useful for Echo Dot to add an extra sentence at the end of the conversation, like “Do you want me to read all your active reminders?” or “You can ask me to repeat all your active reminders”. This could also apply to other cases like timer. Adding a sentence such as “Do you want me to reminder you before the timer stops?” then reading out how much time is left in the timer can potentially reduce the anxiety of the users who consistently search for feedback. This kind of voice feedback allows users to monitor the ongoing activities (Moore, 2013). The researcher also suggests to rethink the use of color on the LED lights. Echo Dot is very careful with the “warning” colors, red and orange. Red and orange colors are used when WiFi is disconnected or the microphone is muted (Martin, 2018). The blue color, as the hallmark color, is the only color used during the interaction. Using a more strong color to convey the incapability of a task or incomprehension of the commands can erase users’ confusion and enhance the feedback. An animation could work as well.

5.1.3 Building Proper User Expectations

It is still unclear what role Echo Dot wants to serve in home setting. Unlike what Luger & Sellen (2016) described in their study, that users hold high expectations of the system intelligence of virtual assistant, participants in this study had relatively low expectations of Echo Dot. Even those participants who saw the anthropomorphism of Echo Dot, only considered those human-like responses as a small aspect of Echo Dot. None of the participants has unrealistic expectations. But the participants experienced what Luger & Sellen (2016) described as “economy of interaction”, that the users stop exploring complex skills of Echo Dot and only engage in those tasks they believe Echo Dot is capable of. It is true that engaging with simple

tasks rises up the task success rate, but “it was ultimately mechanistic and shut down opportunities for developing meaningful CA use” (Luger & Sellen, 2016, p. 5294).

The low expectations of participants have a few reasons. One of the reasons has not been mentioned in previous studies, the influence of the appearance of the device. The relationship between the appearance of Echo Dot and expectations of users still needs empirical research, but from the results of this study, the resemblance to a speaker has influenced user’s mental model of this product. The mental model of a Bluetooth speaker has twisted the way Echo Dot works. Being a Bluetooth speaker means it relies on smartphone to display the content, but the fact is the other way around, the smartphone, more specifically, the Alexa App shows the content from Echo Dot. For example, playing music, one of most engaged activity, is completed by Echo Dot streaming from music resource directly (Amazon, n.d.). Alexa App just displays what the device is streaming. Users do not need to how the system works, but the mental model of Bluetooth speaker can lower the expectations of Echo Dot as Bluetooth speaker works as appendant of smartphone. It is understandable that using the appearance of Bluetooth speaker can enter users’ lives easier, while in the future, Amazon should consider changing the appearance to something that communicates the right intelligence level of Alexa.

Another reason the researcher found that lowered user’s expectations of Echo Dot is the participants’ self-efficacy (Bandura, 1997). Self-efficacy stands for people’s confidence of conducting certain tasks. In this study self-efficacy is the confidence about new technology. A dilemma was found in this study. The participants with high self-efficacy tend to have strong confidence on knowing what Echo Dot is capable of. They believe they know everything of Echo Dot, then they explore very little of the capabilities and stick with what they have already known. The participants who have low self-efficacy tend to stay in their comfort zone, which also limits

the exploring process. This echoes what Vorvoreanu et al. (2019) found in their study, that people with low self-efficacy tend not to use hard-to-use features. In this study, participant 1 tried the translation function of Echo Dot but the results were disappointing, then she blamed herself for not pronouncing correctly and not using the function in a right way. To help people with low-self-efficacy use complex tasks smoothly, Echo Dot needs to make clear how well it can do those tasks. This is one of the human-AI interaction guidelines (Amershi et al., 2019), the clarify of the system performance can "help the user understand how often the AI system may make mistakes" so to avoid users blaming themselves for the mistakes.

The participants have transferred what they learned from friends or past experience with Virtual Assistant like Siri to Echo Dot. This has made the participants' initial use of Echo Dot easier, but also limits users' imagination of the product. Echo Dot entered the market in 2015 (Amazon, n.d.), it is no longer a brand new technology to the users. According to Sciuto et al. (2018), the most frequent Alexa commands are music related, weather inquiring and home automation. These are very simple tasks that can even be completed by a traditional virtual assistant. Amazon needs to differentiate Echo Dot from traditional smartphone-based Virtual Assistant. This requires Amazon and the field to rethink what the HVA is, what kind of assistance it can offer and where the future of HVAs lay (Cohen et al., 2016).

5.1.4 Privacy

Users have always concerned about the data collection of using HVAs especially because the data used by HVAs is not transparent (Cowan, et al., 2017). The privacy concern affects users' adoption of HVAs. People start realizing the trade-off between the personal data and the convenience that can be provided in exchange, but there is still risks that the data can be used in unwanted ways (Cohen, et al., 2016).

Amazon is making efforts to show what data has been collected by showing all the dialogue history in Alexa App, but how the data is used is still unclear. To eliminate users' concern about privacy, Amazon needs to offer as much transparency of the data usage as possible.

For those users who do not want Echo Dot to listen all the time, Amazon should make it clear that users can mute the microphones and unmute only when they use the device. Amazon could provide the option that the microphones only work when a certain button is pressed, it automatically mute the microphones when the button is released.

5.2 Limitations

Due to the timeframe of this study, the participants only had one week to use Echo Dot then were interviewed by the researcher, and the participants are all students from Purdue University, except one participant is from a nearby tech company. The study week conflicted some of the participants' midterm exams, which left them less time using the device. Two of the participants mentioned they were too busy to explore the capabilities of Echo Dot. But fortunately, the participants interacted with Echo Dot much more than the researcher expected, which allowed the researcher to gather meaningful data.

The occupation of student has limited the usage of Echo Dot, as many functionalities such as traffic information and calendar are potentially less useful for students. The student accommodations also limit the choice of Echo Dot's placement. A study facing general public could reveal more information about how the placement influences participant's use of Echo Dot.

The gender was not balanced in the sample. It was not expected that most participants reached out to the researcher were female students. A gender-balanced study could provide more insights in gender-related issues.

The study results differ from some previous studies. Participants of this study did not hold unrealistically high expectations of Echo Dot. This could be because the Echo Dot is no longer a new technology to the participants and this study did not require the participants having no experience with any other virtual assistants. Though the study is not aimed at studying the mental models of the participants, the results indicate that participant's mental models are different from what previous studies have discovered.

5.3 Conclusion

This study aims to understand how users explore the capabilities of HVAs and the factors influence the users' discovery process. The researcher used a study inspired by diary study, combining survey and interview to collect the data. The results show the activities the users engage in and the factors of the HVA device and the users that influence the discovery process. Three themes emerged from the data, they are: The activities participant engaged to learn the capabilities of Echo Dot, the influence of Echo Dot being screen-less and factors of participants that influenced the initial use. Based on the results, the researcher discusses the current issues of the Echo Dot's design and suggests potential improvement for the HVAs.

There are several directions for future research to consider. One direction is to focus on the appearance of HVAs. This study found the low expectations was patricianly due to the appearance. How does resemblance of Bluetooth influence the users' mental models and usage needs to be explored. Another direction of research could explore new ways of providing feedback. The participants in this study generally found the feedback of Echo Dot lame. This issue could exist in all other screen-less virtual assistants. New feedback design can bring better experience to the users. The last direction is the variance of user groups. In this study, the occupation of student has limited the usage and placement of Echo Dot. Discovering the relationship between user group and usage can help improve the design of HVA in the future.

APPENDIX A. QUESTIONS IN THE SURVEYS

QUESTIONS IN SURVEY 1

1. What activities have you done to discover the capabilities of Echo Dot?
2. Why did you engage in those activities?
3. Have you discovered anything from those activities? If you have, what have you discovered?
4. What issues have you encountered using Echo Dot?
5. What do you like or dislike about Echo Dot?
6. Is there anything else you want to add? (optional)

QUESTIONS IN SURVEY 2

1. Have you discovered any new capabilities of Echo Dot?
2. What are the new capabilities you found? How did you find them?
3. What stopped you from discovering new capabilities?
4. What issues have you encountered using Echo Dot?
5. What do you like or dislike about Echo Dot?
6. Will you continue using Echo Dot after the study? Why?
7. Is there anything else you want to add? (optional)

APPENDIX B. INTERVIEW PROTOCOL

DEMOGRAPHIC QUESTIONS

1. What's your major?
2. How old are you?
3. What's your mother language?
4. How familiar with technology do you think you are?

GENERAL QUESTIONS

1. How often do you use virtual assistants before the study?

If the user uses virtual assistants regularly, ask:

- a. What tasks do you engage in with virtual assistants?
- b. Describe one experience with virtual assistants to me.

If the user does not use virtual assistants regularly, ask:

- a. What makes you not using virtual assistants?
- b. What issues with virtual assistants have you encountered?
2. How would you describe the overall experience in the diary study?
3. Will you keep Amazon Echo Dot for the future? Why?
4. Describe a day with Amazon Echo Dot to me.
5. When do you find Amazon Echo Dot most useful?
6. Describe the most pleasing experience with Amazon Echo Dot to me.
7. Describe the most disappointing experience with Amazon Echo Dot to me.

QUESTIONS ABOUT DISCOVERY ACTIVITIES

1. How did you start your discovery process?
2. What activities have you engaged in?
3. What discovery activity do you find most helpful? Why?
4. How would you describe the activities you engaged in with Amazon Echo Dot?

QUESTIONS ABOUT INFLUENCING FACTORS IN DISCOVERY PROCESS

1. What factors do you think influenced you on the discovery process?
2. Where did you set up your Amazon Echo Dot?
3. How motivated do you think you were using Amazon Echo Dot?

QUESTIONS ABOUT ENCOUNTERED ISSUES IN DISCOVERY PROCESS

1. What issues have you encountered?
2. What do you think caused that issue?
3. Did you solve that issue? How did you solve that issue?

REFERENCES

- Allen, J. F., Schubert, L. K., Ferguson, G., Heeman, P., Hwang, C. H., Kato, T., ... Traum, D. R. (1995). The TRAINS project: a case study in building a conversational planning agent. *Journal of Experimental & Theoretical Artificial Intelligence*, 7(1), 7-48.
doi:10.1080/09528139508953799
- Amazon.com Help: View Your Dialog History. (n.d.). Retrieved from
<https://www.amazon.com/gp/help/customer/display.html?nodeId=201602040>
- Amazon. (n.d.). Amazon.com help: Ways to listen to music & media on Alexa. Retrieved from
<https://www.amazon.com/gp/help/customer/display.html?nodeId=201549700>
- Amazon. (n.d.). Echo Dot (3rd generation) | Alexa-enabled bluetooth speaker. Retrieved from
https://www.amazon.com/dp/B0792K2BK6/ref=ods_gw_d_easter19_donut_aucc?pf_rd_p=77d4821b-d4c7-43e0-8d29-be79c1859dd5&pf_rd_r=KS7FV50VHGS40FRZ9ER3
- Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... Horvitz, E. (2019). Guidelines for human-AI interaction. *CHI Conference on Human Factors in Computing Systems*, 13. doi:10.1145/3290605.3300233
- Apple. (n.d.). HomePod. Retrieved from
https://www.apple.com/homepod/?afid=p238%7CsXvn9XQee-dc_mtid_1870765e38482_pcrd_341835648212_&cid=aos-us-kwgo---slid--homepod-e-product-
- Aron, J. (2011, October 26). How innovative is Apple's new voice assistant, Siri? Retrieved from <https://www.newscientist.com/article/mg21228365-300-how-innovative-is-apples-new-voice-assistant-siri/>

- Assefi, M., Liu, G., Wittie, M. P., & Izurieta, C. (2015). An experimental evaluation of apple Siri and google speech recognition. *Proceedings of the 2015 ISCA SEDE*, 1-6.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. London, England: Macmillan.
- Bellegarda, J. R. (2013). Large-scale personal assistant technology deployment: The Siri experience. *Large-scale personal assistant technology deployment: The siri experience*, 2029-2033.
- Berg, B. L., & Lune, H. (2012). *Qualitative Research Methods for the Social Sciences*. Boston, MA: Allyn & Bacon.
- Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., Thompson, H., & Winograd, T. (1977). GUS, a frame-driven dialog system. *Artificial Intelligence*, 8(2), 155-173. doi:10.1016/0004-3702(77)90018-2
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. doi:10.1191/1478088706qp063oa
- Braun, V., & Clarke, V. (2012). Thematic analysis. *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological*, 2, 57-71. doi:10.1037/13620-004
- Carroll, C., Chiodo, C., Lin, A. X., Nidever, M., & Prathipati, J. (2017). Robin: Enabling independence for individuals with cognitive disabilities using voice assistive technology. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*. doi:10.1145/3027063.3049266
- Carter, S., & Mankoff, J. (2005). When participants do the capturing. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '05*. doi:10.1145/1054972.1055098

Cho, J. (2018). Mental models and home virtual assistants (HVAs). *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*.

doi:10.1145/3170427.3180286

Cohen, P., Cheyer, A., Horvitz, E., Kaliouby, R. E., & Whittaker, S. (2016). On the future of personal assistants. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16*, 1032-1037.

doi:10.1145/2851581.2886425

Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About face: the essentials of interaction design*. Hoboken, NJ: John Wiley & Sons.

Corbett, E., & Weber, A. (2016). What can i say? addressing user experience challenges of a mobile voice user interface for accessibility. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '16*. doi:10.1145/2935334.2935386

Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., ... Bandeira, N. (2017). "What can I help you with?": Infrequent users' experiences of intelligent personal assistants. *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '17*.

doi:10.1145/3098279.3098539

Druga, S., Williams, R., Breazeal, C., & Resnick, M. (2017). "Hey Google is it OK if I eat you?". *Proceedings of the 2017 Conference on Interaction Design and Children - IDC '17*. doi:10.1145/3078072.3084330

- Etherington, D. (2015, June 25). Amazon unbundles alexa virtual assistant from echo with new dev tools. Retrieved from <https://techcrunch.com/2015/06/25/amazon-unbundles-alexa-virtual-assistant-from-echo-with-new-dev-tools/>
- Etikan, I. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1. doi:10.11648/j.ajtas.20160501.11
- Ferguson, G., & Allen, J. F. (1998). TRIPS: An integrated intelligent problem-solving assistant. *Proceedings of the National Conference on Artificial Intelligence (AAAI98)*, 567-573. doi:10.1.1.49.8206
- Flaherty, K. (2016, June 5). Diary studies: Understanding long-term user behavior and experiences. Retrieved from <https://www.nngroup.com/articles/diary-studies/>
- Furqan, A., Myers, C., & Zhu, J. (2017). Learnability through adaptive discovery tools in voice user interfaces. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*. doi:10.1145/3027063.3053166
- Følstad, A., Brandtzaeg, P. B., Feltwell, T., Law, E., Tscheligi, M., & Luger, E. A. (2018). Chatbots for social good. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1-4. doi:10.1145/3170427.3185372
- Google. (n.d.). Google Assistant, your own personal Google. Retrieved from https://assistant.google.com/?gclid=Cj0KCQjw-tXIBRDWARIsAGYQAmDrjek_KdmNYqVTcvVCnzofH_dte_r80dgVuLdJN2ZZm6JQITwY7kaAuLWEALw_wcB&gclsrc=aw.ds
- Gorin, A. L., Riccardi, G., & Wright, J. H. (1997). How may I help you? *Speech Communication*, 23(1-2), 113-127. doi:10.1016/S0167-6393(97)00040-X

- Guzzoni, D., Baur, C., & Cheyer, A. (2006). Active: a unified platform for building intelligent web interaction assistants. 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops. doi:10.1109/wi-iatw.2006.27
- Hemphill, C. T., Godfrey, J. J., & Doddington, G. R. (1990). The ATIS spoken language systems pilot corpus. Proceedings of the workshop on Speech and Natural Language - HLT '90. doi:10.3115/116580.116613
- Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Medical Reference Services Quarterly*, 37(1), 81-88. doi:10.1080/02763869.2018.1404391
- Jadeja, M., & Varia, N. (2017). Perspectives for evaluating conversational AI. *CEUR Workshop Proceedings*, 2065(6), 54-57. doi:10.1145/nnnnnnn.nnnnnnn
- Janarthanam, S. (2017). *Hands-on chatbots and conversational UI development: Build chatbots and voice user interfaces with chatfuel, dialogflow, Microsoft bot framework, Twilio, and Alexa skills*.
- Kearney, A. R., & Kaplan, S. (1997). Toward a methodology for the measurement of knowledge structures of ordinary people. *Environment and Behavior*, 29(5), 579-617.
doi:10.1177/0013916597295001
- Kinsella, B. (2018, September 2). Amazon Alexa now has 50,000 skills worldwide, works with 20,000 devices, used by 3,500 brands. Retrieved from <https://voicebot.ai/2018/09/02/amazon-alexa-now-has-50000-skills-worldwide-is-on-20000-devices-used-by-3500-brands/>
- Kinsella, B. (2017, December 5). Alexa skill count grows 300% over the past year in u.s. but growth rate declining. Retrieved from <https://voicebot.ai/2017/12/05/alexa-skill-count-grows-300-over-the-past-year-in-u-s-but-growth-rate-declining/>

- Kinsella, B. (2018, March 22). Amazon alexa skill count surpasses 30,000 in the u.s. Retrieved from <https://voicebot.ai/2018/03/22/amazon-alexa-skill-count-surpasses-30000-u-s/>
- Kinsella, B. (2018, April 2). Smart speaker owners use voice assistants nearly 3 times per day. Retrieved from <https://voicebot.ai/2018/04/02/smart-speaker-owners-use-voice-assistants-nearly-3-times-per-day/>
- Koetsier, J. (2018, August 2). Amazon echo, google home installed base hits 50 million; apple has 6% market share, report says. Retrieved from <https://www.forbes.com/sites/johnkoetsier/2018/08/02/amazon-echo-google-home-installed-base-hits-50-million-apple-has-6-market-share-report-says/#4524167b769c>
- Koetsier, J. (2018, May 31). Smart speaker users growing 48% annually, to hit 90m in usa this year. Retrieved from <https://www.forbes.com/sites/johnkoetsier/2018/05/29/smart-speaker-users-growing-48-annually-will-outnumber-wearable-tech-users-this-year/#50ec09ec5dde>
- Kopp, S., Gesellensetter, L., Krämer, N. C., & Wachsmuth, I. (2005). A conversational agent as museum guide – Design and evaluation of a real-world application. *Intelligent Virtual Agents*, 329-343. doi:10.1007/11550617_28
- Lei, X., Tu, G., Liu, A. X., Li, C., & Xie, T. (2018). The insecurity of home digital voice assistants - Amazon Alexa as a case study. *2018 IEEE Conference on Communications and Network Security (CNS)*, 1-13. doi:10.1109/cns.2018.8433167
- Luger, E., & Sellen, A. (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*. doi:10.1145/2858036.2858288

- López, G., Quesada, L., & Guerrero, L. A. (2018). Alexa vs. Siri vs. Cortana vs. Google Assistant: a comparison of speech-based natural user interfaces. *Advances in Intelligent Systems and Computing*, 241-250. doi:10.1007/978-3-319-60366-7_23
- Martin, T. (2018, December 21). What all those lights on your Amazon Echo mean. Retrieved from <https://www.cnet.com/how-to/light-ring-colors-amazon-echo-alexa/>
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach*. Thousand Oaks: SAGE Publications.
- McGregor, M., & Tang, J. C. (2017). More to meetings: Challenges in using speech-based technology to support meetings. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. doi:10.1145/2998181.2998335
- McTear, M., Callejas, Z., & Griol, D. (2016). *The conversational interface: Talking to smart devices*.
- Microsoft. (n.d.). Personal digital assistant - Cortana home assistant - Microsoft. Retrieved from <https://www.microsoft.com/en-us/cortana>
- Moore, R. K. (2013). Spoken language processing: Where do we go from here? *Lnai*, 7407, 119-133. doi:10.1007/978-3-642-37346-6_10
- Moore, R. K. (2016). Is spoken language all-or-nothing? Implications for future speech-based human-machine interaction. *Lecture Notes in Electrical Engineering*, 281-291. doi:10.1007/978-981-10-2585-3_22
- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI '94*. doi:10.1145/191666.191729

- Norman, D. (1983). "Some observations on mental models". *Mental Models*, 7-14.
- Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. New York: Basic Books.
- Palen, L., & Salzman, M. (2002). Voice-mail diary studies for naturalistic data capture under mobile conditions. *Proceedings of the 2002 ACM conference on Computer supported cooperative work - CSCW '02*. doi:10.1145/587091.587092
- Patton, J. Q. (2015). *Qualitative research design*. Thousand Oaks, CA: SAGE.
- Pearl, C. (2017). *Designing voice user interfaces*. US: O'Reilly media.
- Porcheron, M., Fischer, J., Reeves, S., & Sharples, S. (2018). Voice interfaces in everyday life. *Chi*, 1-12. Retrieved from doi.org/10.1145/3173574.3174214
- Porcheron, M., Fischer, J. E., & Sharples, S. (2017). Do animals have accents?": Talking with agents in multi-party conversation. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. doi:10.1145/2998181.2998298
- Portela, M., & Granell-Canut, C. (2017). A new friend in our smartphone? Proceedings of the XVIII International Conference on Human Computer Interaction - Interacción '17. doi:10.1145/3123818.3123826
- Portet, F., Vacher, M., Golanski, C., Roux, C., & Meillon, B. (2011). Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects. *Personal and Ubiquitous Computing*, 17(1), 127-144. doi:10.1007/s00779-011-0470-5
- Pradhan, A., Mehta, K., & Findlater, L. (2018). "Accessibility came by accident": Use of voice-controlled intelligent personal assistants by people with disabilities. *Proceedings of the*

2018 CHI Conference on Human Factors in Computing Systems - CHI '18.

doi:10.1145/3173574.3174033

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). “Alexa is my new bff”: social roles, user satisfaction, and personification. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*. doi:10.1145/3027063.3053246

Rabiner, L. R., & Levinson, S. E. (1990). Isolated and connected word recognition— Theory and selected applications. *Readings in Speech Recognition*, 115-153. doi:10.1016/b978-0-08-051584-7.50014-0

Reeves, S. (2017). Some conversational challenges of talking with machines. *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17 Companion*, 431-436. doi:10.1145/3022198.3022666

Reis, A., Paulino, D., Paredes, H., & Barroso, J. (2017). Using intelligent personal assistants to strengthen the elderlies’ social bonds. *Universal Access in Human–Computer Interaction. Human and Technological Environments*, 593-602. doi:10.1007/978-3-319-58700-4_48

Reis, A., Paulino, D., Paredes, H., & Barroso, J. (2017). Using intelligent personal assistants to strengthen the elderlies’ social bonds. *Universal Access in Human–Computer Interaction. Human and Technological Environments*, 593-602. doi:10.1007/978-3-319-58700-4_48

Rong, X., Fourney, A., Brewer, R. N., Morris, M. R., & Bennett, P. N. (2017). Managing uncertainty in time expressions for virtual assistants. *Proceedings of the 2017 CHI*

Conference on Human Factors in Computing Systems - CHI '17.

doi:10.1145/3025453.3025674

Rosenfeld, R., Olsen, D., & Rudnicky, A. (2001). Universal speech

interfaces. *interactions*, 8(6), 34-44. doi:10.1145/384076.384085

Rudzicz, F., Wang, R., Begum, M., & Mihailidis, A. (2015). Speech interaction with personal

assistive robots supporting aging at home for individuals with Alzheimer's disease. *ACM Transactions on Accessible Computing*, 7(2), 1-22. doi:10.1145/2744206

Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018). "Hey Alexa, what's up?": Studies of in-

home conversational agent usage. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*. doi:10.1145/3196709.3196772

Shawar, B. A., & Atwell, E. (2007). Chatbots: Are they really useful? *LDV-Forum*.

Sohn, T., Li, K. A., Griswold, W. G., & Hollan, J. D. (2008). A diary study of mobile

information needs. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. doi:10.1145/1357054.1357125

Thurman, J. (2008, June 22). Kubrick's Frankenstein: HAL in 2001: A Space Odyssey.

Retrieved from <https://cineprism.wordpress.com/2007/12/11/kubrick's-frankenstein-hal-in-2001-a-space-odyssey/>

Trappl, R. (2013). Introduction: from Jeeves to Jeannie to Siri, and then? Your Virtual Butler, 1-

8. doi:10.1007/978-3-642-37346-6_1

Traum, D. R., & Larsson, S. (2003). *The information state approach to dialog management*.

Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., & Bickmore, T. (2012). Designing

relational agents as long term social companions for older adults. *Intelligent Virtual Agents*, 289-302. doi:10.1007/978-3-642-33197-8_30

- Vassallo, G., Pilato, G., Augello, A., & Gaglio, S. (2010). Phase coherence in conceptual spaces for conversational agents. *Semantic Computing*, 357-371.
doi:10.1002/9780470588222.ch18
- Vorvoreanu, M., Zhang, L., Huang, Y., Hilderbrand, C., Steine-Hanson, Z., & Burnett, M. (2019). From gender biases to gender-inclusive design: An empirical investigation. CHI Conference on Human Factors in Computing Systems Proceedings, 53. Retrieved from <https://doi.org/10.1145/3290605.3300283>
- Weizenbaum, J. (1966). ELIZA---a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.
doi:10.1145/365153.365168
- Whitenton, K. (2016, January 31). Voice interaction UX: Brave new world...same old story. Retrieved from <https://www.nngroup.com/articles/voice-interaction-ux/>
- Whitenton, K., & Budiu, R. (2018, September 16). Voice interaction UX: Brave new world...same old story. Retrieved from <https://www.nngroup.com/articles/voice-interaction-ux/>
- Wilpon, J., DeMarco, D., & Mikkilineni, R. (1988). Isolated word recognition over the DDD telephone network. Results of two extensive field studies. ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing.
doi:10.1109/icassp.1988.196508
- Wilpon, J., Rabiner, L., Lee, C., & Goldman, E. (1990). Automatic recognition of keywords in unconstrained speech using hidden Markov models. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(11), 1870-1878. doi:10.1109/29.103088

Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3(1), 1-191.

doi:10.1016/0010-0285(72)90002-3

Wolters, M. K., Kelly, F., & Kilgour, J. (2016). Designing a spoken dialogue interface to an intelligent cognitive assistant for people with dementia. *Health Informatics Journal*, 22(4), 854-866. doi:10.1177/1460458215593329

doi:10.1177/1460458215593329

Yankelovich, N. (1996). How do users know what to say? *interactions*, 3(6), 32-43.

doi:10.1145/242485.242500