Exploring User Experience with Voice Assistants: Impact of Prior Experience on Voice Assistants.

Lawal Ibrahim Dutsinma Faruk
School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
lawal.faruk@mail.kmutt.ac.th

Suree Funilkul
School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
suree@sit.kmutt.ac.th

Pornchai Mongkolnam
School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
pornchai@sit.kmutt.ac.th

Prateep Puengwattanapong
School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
ppuengwattanapong@gmail.com

Debajyoti Pal
Innovative Cognitive Computing Research Center (IC2), School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
debajyoti.pal@mail.kmutt.ac.th

ABSTRACT

Voice assistants (VA) like Siri, Alexa, Cortana, and Google Assistant are on the rise, and are currently integrated into smartphones, and dedicated home speakers. They handle various tasks through voice commands, from home automation, emails to calendars.in general we can say the VA change how we interact with technology, hence benefiting diverse users. To reap more benefit of the VA, It is crucial to emphasize user-centric research as a focus, in addition to the technical advancements especially given the abundance of commercial VA. Each VA is unique and different, and the option of which one to acquire still remains a topic for discussion. However, prior experience affect the use of technology but if that also affect VA is still yet uncovered. This study aim to uncover how prior experience affect the user experience while using multiple VA, which ultimately affect their overall preference. Understanding VA user experience is crucial as they integrate further into our lives. Researchers and Manufacturers must consider user preferences for broader adoption, hence this study reveal more insights into how VA cater to both experienced and non-experienced users (First timers).

CCS CONCEPTS

• **Human-centered computing** → Human computer interaction (HCI); HCI design and evaluation methods; Usability testing.

KEYWORDS

Voice Assistant, User Experience, Alexa, Google Assistant, Prior Experience, Novice

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IAIT 2023, December 06–09, 2023, Bangkok, Thailand

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0849-7/23/12…$15.00
https://doi.org/10.1145/3628454.3629470

ACM Reference Format:
Lawal Ibrahim Dutsinma Faruk, Suree Funilkul, Pornchai Mongkolnam, Prateep Puengwattanapong, and Debajyoti Pal. 2023. Exploring User Experience with Voice Assistants: Impact of Prior Experience on Voice Assistants.. In *13th International Conference on Advances in Information Technology (IAIT 2023), December 06–09, 2023, Bangkok, Thailand*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3628454.3629470

1 INTRODUCTION

Voice assistants (VA) are intelligent computer programs that are capable of understanding and responding to human speech through synthesized voices, with some of the most popular being Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Assistant. VA are currently integrated into as smartphones, cars, as well as dedicated smart speakers. While interacting with the VA, users are expected to speak out voice commands which enable then allow the VA perform various tasks; These tasks include asking questions, managing home automation devices [34], controlling media playback [2], carrying out essential functions like handling emails, managing to-do lists, and scheduling appointments [5] and so on. Moreover, the VA have demonstrated their significant impact in various areas in human lives, including but not limited to education [26, 27], aiding users who struggle with the use of Graphical User Interfaces (GUI) system [3], and enhancing social bonds among the elderly [25]. In fact It is safe to say the VA have revolutionize the traditional mode of interaction with systems, by omitting the use and dependency of the GUI [6].

There has been a lot of buzz towards enhancing the machine capabilities of VA [12, 28], which signifies efforts to improve their technical aspects such as level of accuracy and intelligence. However, while there is no doubt this has greatly advanced their capability, it is also equally important to emphasize on the VA user-centric aspect, in order to promote their widespread adoption [32]. Ultimately, investigating the user experience (UX) of VA will improve their user perception, temper unrealistic expectations [14], and ultimately speed up their acceptance [10]. The UX of emergent technologies such as conversational agents varies on their capabilities [4]. On that aim, commercial VA are designed to effectively

execute a broad array of tasks capabilities. Nevertheless, their distinctive capabilities, intonation, personalities, and features set them apart from one another depending on the manufacturers, hence potentially leading to varying preferences among different users [6]. Numerous commercial VA are available in the market, with Amazon's Alexa and Google Assistant emerging as prominent contenders [11]. While both are designed to perform similar tasks, each device exhibits unique qualities in terms of personality, task execution, voice tones, and accessories. For instance, Google Assistant, which is established more than a decade ago, seamlessly offers voice-controlled access to Google services [7]. In contrast, Alexa distinguishes itself with its user-friendly interface [21] and also comes with easier accessibility to Amazon services. Ultimately, the choice between Alexa and Google Assistant depends on the individual preferences, since their core functionality are similar. Nevertheless, it is also noteworthy previous studies have shown acceptance of technology devices can be influenced not only by individual personal preference, or user expectation, but also by prior experience.

User prior experience has been shown to influence technology usage and acceptance. For instance [29] highlights experienced users exhibit more confidence level and ease of use while using technologies, which increases usage satisfaction and ultimately adoption. Consequently, user prior experience plays a pivotal role in shaping perceptions of technology usefulness, with experienced users more likely to find new technologies valuable and effective. In contrast, non-experienced users may hold lower expectations due to a lack of understanding of a product, which result in difficulty during use, potentially leading to frustration and abandonment [15].

Several studies have delved into the user experience of Google Assistant and Alexa concerning their functionality [9], and their personality [6, 20]. However, the uncharted territory lies in understanding how does the user prior experience affect the UX while using VA, and whether prior experience impact the VA ease of use, performance, sentimentality and overall satisfaction. Therefore this study tend to uncover if the VA UX differs between experienced and non-experienced users of VA while utilizing both Amazon Alexa and Google Nest which is a dedicated Google Assistant speaker. Based on our study objective, we propose the following hypothesis:

H1*:* Experienced users, in comparison to non-experienced users, will exhibit an overall enhanced user experience when interacting with the voice assistant, as reflected in their ease of use, satisfaction, positive sentiments, and VA useful performance.

Since the VA represent a unique mode of interaction, differing from traditional GUI-based interfaces. The debate surrounding the effect of their human-like conversational abilities and whether they evoke positive or negative reactions, which as well also include concerns related to the uncanny valley is still ongoing [31]. However, based on users' prior experience, we anticipate they will be familiar with it more than the non-experienced users. Therefore, we propose the following hypothesis:

H2: Users with prior experience will have higher approval levels with the voice assistant's human-like conversational mannerisms compared to users with no prior experience.

The benefit of our study include; VA developers stand to gain valuable insights into user needs, with provision of opportunity to create a more user-friendly VA that cater to a diverse user group. Additionally, from a marketing perspective, by examining the UX interactions of both user groups (experienced and non-experienced users) with both VA, the discoveries can guide VA marketers in creating messages that resonate with specific user group. Moreover, the study has the potential to highlight the unique strengths and weaknesses UX aspect of Google Nest and Alexa while used by both user groups (experienced and non-experienced users), which can inform product development efforts, and ultimately drive adoption and enhance user satisfaction.

2 RECENT WORK

This section delve into studies that has been conducted to assess the effect of prior experience in today's emergent technology. Given that the main focus of our study is the comparative analysis between experienced and non-experienced users while interacting with VA. This section further explores studies that were carried out to provide more perspectives to our study objectives.

A User's Prior experience is shown to affect the user behavior and UX when it comes to technology usage. A study by [30] found that users with prior experience developed a set of expectations and beliefs about the technology, which influence their perceptions of its usefulness and ease of use. Additionally, prior experience can lead to greater confidence in one's ability to use the technology, which increase behavioral control, and ultimately lead to a better usage. Another study conducted by [22] highlights the impact of users' previous experiences on gameful interaction design within information and communication technology (ICT) for the elderly. The study emphasizes that if users have encountered prior negative experiences, such as challenges in learning and using technology, or if they have no prior experience at all, it can evoke feelings of fear and helplessness when engaging with the technologies. Prior experience have also been shown to directly affect more sophisticated technology. A study by [24] found that prior experience with self-driving cars had a significant effect on acceptance. Respondents with previous experience showed higher acceptance than respondents without such experiences. Familiarity with autonomous vehicles can reduce worries and have a positive effect on attitudes toward them, moreover drivers with prior knowledge of autonomous vehicles are more willing to relinquish driving control. Moreover, user prior experience also affect the user buoyancy and also increase user adoption [19]. On that aim, [33] demonstrate users who have prior experience with technology tools have better chances to adapt and use them. Furthermore, users with prior experience are better able to take advantage of new technologies. In the realm of conversational agents, [13] suggests prior experience with chatbot is an important variable in determining their acceptance. Specifically, the authors suggest that previous use of chatbot impacts their perceived ease of use and perceived usefulness, which, in turn, can influence behavioral intention to use them.

In summary, there exists a well-documented understanding that users prior experience serves as a critical factor that influence the positive technology usage, UX, and user acceptance. However, there remains a lack of research dedicated to assessing how user prior experience applied to VA UX. Addressing this research gap is vital

Table 1: Dimensions and Item list.

User Experience Dimension	Item Selected	Scale Item Originated from
Performance refer to the VA, reliability, speed, and accuracy in achieving user goals and desired outcomes.	The system is accurate.	SASSI
	I felt confident using this system.	Speech User Interface Service Quality(SUISQ)
	I was able to complete the tasks and scenarios quickly using this system.	The PSSUQ Survey
	The voice-assistant was unreliable.	SASSI
	The voice-assistant didn't had difficulty in understanding what I asked it to do.	Voice Usability Scale (VUS)
Ease of use refers to the simplicity and intuitiveness of a VA, making it easy for users to understand, navigate, and interact with.	The system is easy to use.	SASSI
	It was easy to use voice to perform the tasks in this system.	User Experience Evaluation of Conversational Systems(UEXECS)
	It is clear how to speak to the system.	SASSI
	A high level of concentration is required when using the system.	SASSI
	I found it easy to understand how to interact by voice in the system.	User Experience Evaluation of Conversational Systems(UEXECS)
	I could find what I needed without any difficulty.	Speech User Interface Service Quality(SUISQ)
Satisfaction refer to the level of contentment and fulfillment a user feels after interacting with a VA.	Overall, I am satisfied with using the voice-assistant.	Voice Usability Scale (VUS)
	I would likely use this system again.	Speech User Interface Service Quality(SUISQ)
	The system behaved the way I expected during the voice interaction.	User Experience Evaluation of Conversational Systems(UEXECS)
Conversation Mannerism refers to the way in which a VA communicates with users through natural language and conversation.	The system seemed courteous.	Speech User Interface Service Quality(SUISQ)
	The system seemed professional in its speaking style.	Speech User Interface Service Quality(SUISQ)
	The system voice sounded enthusiastic of full energy.	Speech User Interface Service Quality(SUISQ)
	The system seemed friendly.	Speech User Interface Service Quality(SUISQ)
	The system seemed polite.	Speech User Interface Service Quality(SUISQ)
Sentiment refers to the emotional reactions or feelings that users experience when interacting with a VA.	I consider the system a pleasant conversational partner.	Personality, Usability, and Enjoyability of Voice Agents(PUEVA)
	I had fun using my voice to perform these tasks in the system.	User Experience Evaluation of Conversational Systems(UEXECS)
	The interaction with the system is boring.	SASSI
	The system makes me happy when I talk with it.	Personality, Usability, and Enjoyability of Voice Agents(PUEVA)
	I enjoyed using the system.	SASSI

for a more comprehensive understanding of UX, user behavior and preferences in the rapidly evolving landscape of VA technology.

3 METHODOLOGY

This section include the presentation of the scales we used to measure the UX of VA, the formulation of task, and the development of experimental designs. By integrating these critical components, our methodology seeks to facilitate a deeper understanding of how the UX dimensions relate to both user group (experienced and non-experienced) when it comes to VA.

3.1 User experience dimension

The first step in the process is selecting the items we used to assess the UX dimension in our research. To accomplish this, we carried out a literature review on previously developed scale for VA, aligning them with the specific UX dimension we aimed to measure. The process commenced by identifying the dimensions, encompassing a definition of these dimensions. Subsequently, we conducted a comprehensive review of the items from various scales, and selected the items that conceptually matched our proposed dimensions. Our scale was made up of total 24 items representing 5 dimensions, thus each item denoted by a 5-point Likert scale, with "Strongly Disagree" (1) as the lowest and "Strongly Agree" (5) as the highest rating. Table 2 represent the items we utilize and the scale they originated from.

Performance dimension represent the speed, reliability and accuracy in carrying out task, we selected items from the PSSUQ (Post-Study System Usability Questionnaire) scale that assess system

Table 2: Interactive Tasks and Commands

Categories	Task	Commands
Productive task. Tasks that aid users in achieving meaningful objectives enhances their efficiency, goal attainment, and overall satisfaction.	Setting Alarms.	Set an alarm for 10:30 PM.
	Creating To do List.	Create a to-do list.
Entertainment and Media task. Tasks that capture the attention and interest of an audience, provide pleasure and enjoyment.	Jokes.	Tell me a joke
	Playing music.	Play bad blood by Taylor swift.
Travel and Navigation task. Tasks that relate to travel and transit, including hotel recommendations, flight bookings, seeking transit directions.	Flight enquiry.	How long is the flight from London to Bangkok?
	Driving Navigation.	What is quickest route to victory monument?
Health and Wellness task. Tasks that aid in both physical and mental well-being.	Healthy food.	What is a healthy alternative to a burger?
	Meditation and Stress management.	Alexa, play white noise from Spotify.
General Knowledge and Information task. Tasks that enable users to perform internet searches using voice commands such as weather flash and news briefings.	Weather updates.	Will I need an umbrella today in Bangkok?

response time [16]. These item represent the performance dimension because it measure the effectiveness and speed of completing the task without lagging. Similarly, for ease of use dimension, we selected items from scales such as User Experience Evaluation of Conversational Systems (UEXECS) [8]. These items are selected to represent the dimension because they measure the simplicity and intuitiveness of a VA, highlighting how easy it is for users to understand, navigate, and interact with. For sentiment UX dimension, we selected items such as items from the PUEVA (Personality, Usability, and Enjoyability of Voice Agents) that assess the likability and approachability of a VA [18]. Conversation mannerism refers to the speech style and manner in which a VA communicates with users through natural language and conversation, which gives the VA a sense of mental character in the mind of the user. In this dimension we utilized items from Speech User Interface Service Quality (SUISQ) [17]. Sentiment dimension refers to the emotion that's stir up within the user during the interaction with the VA, and we utilized items from scales such as Personality, Usability, and Enjoyability of Voice Agents (PUEVA) [18]. All the items used in our studies and the dimension they measure are presented in Table 1.

3.2 Task

The current stage of commercial VA is still in its early phases, nevertheless, their popularity has been steadily increasing, and new features continue to be added. However, studies have indicates that people predominantly utilize VA for a limited range of tasks, such as playing music, checking the weather, and setting timers [1, 2, 23]. Hence, we utilized and compiled a set of tasks that mirror prevalent real-world use of VA, establishing a fundamental framework compatible with all VA. Our primary objective is to facilitate user interactions with the VA through these tasks. We systematically categorized the tasks according to their intended purpose, each task being associated with specific commands to be spoken to the VA. Table 2 illustrates the task categories along with sample commands within each category.

3.3 Experiment design

3.3.1 Experiment participant. The study was conducted in Thailand, it included 24 participants with a gender distribution of 6 females (25%) and 18 males (75%). The participants' ages ranged from 23 to 38 years, with a mean age of 27. Half of the participants had prior experience and actively used VA (at least twice a week) for the past six months (50%), while the other half had no prior experience (50%). Moreover, all participants had an intermediate level of English proficiency.

3.3.2 Experiment protocol. Prior to commencing the experiment, we conducted a pre-experiment in which we ask about their demographic information and inquired about participants' prior experiences with VA. Afterward we carefully selected 12 participants who have no previous experience with VA, and selected 12 participants who have prior experience. After choosing 24 participant in total we went ahead to carry out the experiment. During the experiment six of the non-experienced participants were selected at random to start their interaction with Amazon Alexa first and then proceeded to interact with Google Nest second, while the other six non-experienced users began with Google Nest first and subsequently used Alexa afterwards. The same usage pattern was applied to the experienced users. This usage pattern was carried because out to avoid bias, in which one VA experience will not be affected base on bias preferential experience. During the interactions, participants were assigned a carry out the task on Table 2 by voicing the commands to engage with the VA. After each interaction with a VA, participants will fill in the items questionnaire to measure their experience with that specific VA. Only then the user

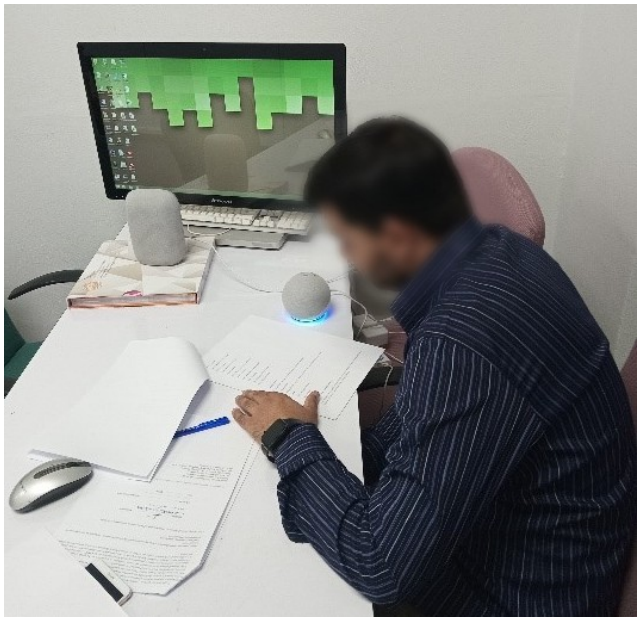

Figure 1: Participant interacting with a Voice Assistant

will interact with the second VA, which also they are told to fill the questionnaire item the second time. Figure 1 shows a participants interacting with a VA during the experiment.

4 RESULT

In this section, we carry out a careful analysis of the data collected during the methodology phase of our study. This phase constitutes as the foundation upon which our findings are constructed, which enable us to extract valuable insights and draw comprehensive conclusions. Initially, following the collection of user evaluations regarding their interactions with VA, we converted their responses into their numerical values ("Strongly Disagree" = 1, "Disagree" = 2, "Neutral" = 3, "Agree" = 4, "Strongly Agree" = 5). Subsequently, we utilized SPSS to transform the collected user data points into variables. We then conducted a descriptive analysis to assess the impact of user prior experience based on the mean averages of each user group while using each VA. As graphically presented in Figure 2, as well as reported in Table 3 and 4, we performed and analyze the independent t-test, also known as the two-sample t-test, to determine if there is a significant difference between the experience of the two independent user groups (experienced and non-experienced users).

The results of an independent sample t-test were employed to explore potential differences between experienced and non-experienced user's base on the five different dimensions of UX (UX) related to Alexa and Google Nest Respectively. These dimensions included "Performance," "Ease of Use," "Satisfaction," "Conversation Mannerism," and "Sentiment." In the case of both user groups experience with Alexa VA as presented on Table 3; Performance dimension, there was no statistically significant difference in UX scores between experienced (M = 3.5167, SD = 0.59365) and non-experienced users (M = 3.8500, SD = 0.80510), as indicated by a non-significant Levene's test for equality of variances (F = 1.403, p = 0.249) and a non-significant t-test (t = -1.154, p = 0.261). Similarly, in the "Ease of Use," "Satisfaction," and "Sentiment" dimensions, there were no significant differences between experienced and non-experienced users. Levene's test for equality of variances showed no significant differences, and the t-tests were non-significant, with confidence intervals that included zero.

However, in the "Conversation Mannerism" dimension, the Levene's test indicated unequal variances (F = 16.795, p = 0.0004). Despite this, the t-test was non-significant (t = -1.921, p = 0.075), but the confidence interval (CI [-0.82210, 0.04433]) was somewhat close to zero, suggesting that there might be a subtle difference between the two groups in this dimension, but it did not reach statistical significance. Therefore base on the experience of both user groups (experienced and non-experienced) using Alexa, the null hypothesis is accepted which is there is no significant difference between the two user groups, which means our H1 hypothesis is rejected when it comes to the Alexa VA. However, there may be a difference in the effect of conversational mannerism, even though it is not significant different which makes H2 hypothesis partially true.

In the case of both user groups experienced with Google Nest VA as presented on Table 4, 'Performance' dimension, experienced users (M = 4.1000, SD = 0.71095) had a significantly higher mean score compared to non-experienced users (M = 3.5167, SD = 0.56862). The assumption of equal variances was met (F = 0.337, p = 0.567), and the t-test was statistically significant (t = 2.220, p = 0.037, 95% CI [0.03831, 1.12835]), indicating that experienced users believe the VA performed significantly better. For the 'Ease of Use' dimension, experienced users (M = 3.8611, SD = 0.31649) also had a higher mean score than non-experienced users (M = 3.5139, SD = 0.41107). The assumption of equal variances held (F = 0.170, p = 0.685), and the t-test was statistically significant (t = 2.318, p = 0.030, 95% CI [0.03663, 0.65781]), signifying that experienced users found the Google Nest VA easier to use. In the 'Satisfaction' dimension, while experienced users (M = 4.0000, SD = 0.76541) had a slightly higher mean score than non-experienced users (M = 3.6111, SD = 0.61682), the assumption of equal variances was met (F = 2.011, p = 0.170), and the t-test was not statistically significant (t = 1.370, p = 0.184, 95% CI [-0.19962, 0.97740]), suggesting no significant difference in satisfaction between the two groups. For the 'Conversational Mannerism' dimension, experienced users (M = 4.1667, SD = 0.47140) had a higher mean score than non-experienced users (M = 3.7917, SD = 0.29409). Although the assumption of equal variances was met (F = 3.219, p = 0.087), the t-test was statistically significant (t = 2.338, p = 0.029, 95% CI [0.04236, 0.70764]), indicating that experienced users exhibited more favorable conversational mannerisms. In the 'Sentiment' dimension, experienced users (M = 3.6667, SD = 0.35887) had a higher mean score compared to non-experienced users (M = 3.2083, SD = 0.43736). The assumption of equal variances was met (F = 1.375, p = 0.253), and the t-test was statistically significant (t = 2.806, p = 0.010, 95% CI [0.11963, 0.79704]), signifying that experienced users expressed more positive sentiments. This goes against the null hypothesis that there is no significant difference between the users with prior experience and user with non-experience. This support our H1 and H2 hypotheses, however in the case of satisfaction dimension there was no significant different between the two user group.

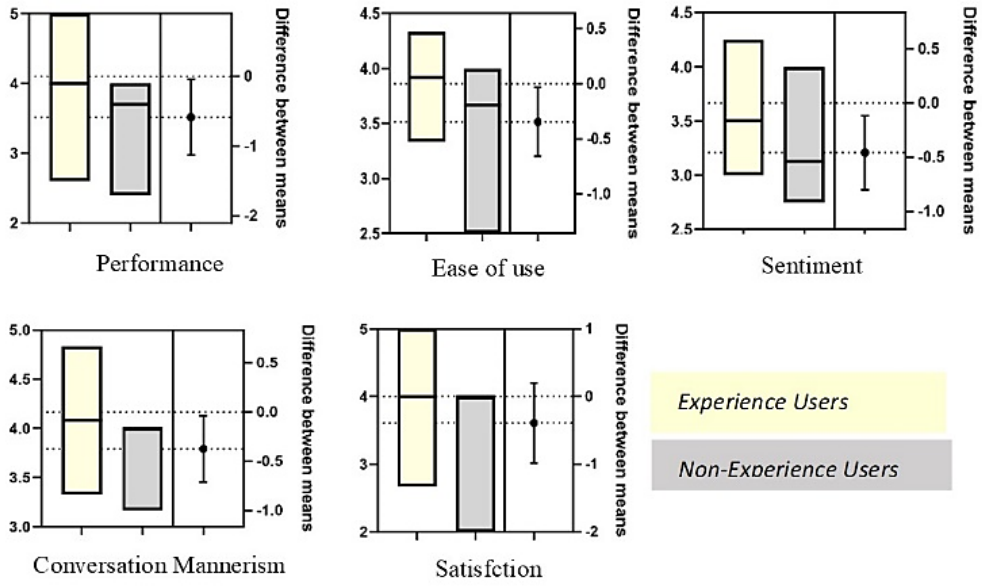

Figure 2: Comparison between experienced users and non-experienced users interacting with Google Nest

Table 3: Independent t-test result for experienced and non-experienced users utilizing Alexa voice assistant.

User Experienced Dimension	User Group	Mean	Std. Deviation	Levene's Test for Equality of Variances		t-test for Equality of Means			
				F	Sig.	t	Sig.(2-tailed)	95% Confidence Interval of the Difference	
								Lower	Upper
Performance (Alexa)	Experienced	3.5167	0.59365	1.403	0.249	-1.154	0.261	-0.93219	0.26552
	Non-Experienced	3.8500	0.80510						
Ease of use (Alexa)	Experienced	3.5694	0.38572	1.020	0.323	-0.317	0.755	-0.52429	0.38540
	Non-Experienced	3.6389	0.65456						
Satisfaction (Alexa)	Experienced	3.5278	0.54045	1.391	0.251	-1.324	0.199	-1.14052	0.25163
	Non Experienced	3.9722	1.02945						
Conversational mannerism (Alexa)	Experienced	3.8333	0.25624	16.795	0.000	-1.921	0.075	-0.82210	0.04433
	Non Experienced	4.2222	0.65263						
Sentiment (Alexa)	Experienced	3.3958	0.47023	0.635	0.434	-0.577	0.570	-0.47879	0.27046
	Non-Experienced	3.5000	0.41286						

In summary, the results indicate that experienced users of Google Nest VA outperformed non-experienced users in the 'Performance,' 'Ease of Use,' 'Conversational Mannerism,' and 'Sentiment' dimensions, while there was no significant difference in the 'Satisfaction' dimension. These findings confirms part of our hypothesis H1 with ease of use, positive sentiments, and perceived VA useful performance considered higher by experienced users , however satisfaction dimension did not support our hypothesis. There was no significant difference in satisfaction between experienced and non-experienced users. Moreover, the finding also support our H2 hypothesis in which users with prior experience approved the VA human-like conversational mannerisms compared to users with no

Table 4: Independent t-test result for experienced and non-experienced users utilizing Google Nest voice assistant.

User Experience Dimension	User Group	Mean	Std. Deviation	Levene's Test for Equality of Variances		t-test for Equality of Means			
				F	Sig.	t	Sig. (2-tailed)	95% Confidence Interval of the Difference	
								Lower	Upper
Performance (Google Nest)	Experienced	4.1000	0.71095	0.337	0.567	2.220	0.037	0.03831	1.12835
	Non-Experienced	3.5167	0.56862						
Ease of use (Google Nest)	Experienced	3.8611	0.31649	0.170	0.685	2.318	0.030	0.03663	0.65781
	Non-Experienced	3.5139	0.41107						
Satisfaction (Google Nest)	Experienced	4.0000	0.76541	2.011	0.170	1.370	0.184	-0.19962	0.97740
	Non-Experienced	3.6111	0.61682						
Conversational mannerism (Google Nest)	Experienced	4.1667	0.47140	3.219	0.087	2.338	0.029	0.04236	0.70764
	Non Experienced	3.7917	0.29409						
Sentiment (Google Nest)	Experienced	3.6667	0.35887	1.375	0.253	2.806	0.010	0.11963	0.79704
	Non Experienced	3.2083	0.43736						

prior experience. More detailed explanation of the result finding is deliberate in the discussion section.

5 DISCUSSION

The discoveries of this study present a nuanced perspective on the role of prior experience in shaping the UX with VA, specifically in Amazon Alexa and Google Nest. While the statistical results have been presented, the results gotten from Amazon Alexa and Google Nest have either oppose or support our hypothesis, it is important to explore the deeper implications and possible aspects contributing to the observed results. Now to delve deeper into our hypothesis, with the first hypothesis which states;

H1: Experienced users, in comparison to non-experienced users, will exhibit an overall enhanced user experienced when interacting with the voice assistance, as reflected in their perceived ease of use, satisfaction, positive sentiments, and perceived VA useful performance.

Whilst using Alexa VA, ease of use, performance, satisfaction, and sentiment dimensions were not significant different between experienced and non-experienced users, which oppose our H1 hypothesis. However. In contrast to Amazon Alexa, Google Nest VA ease of use, performance, and sentiment dimensions were significant different between experienced and non-experienced users, in favor of the experienced users. This support our H1, except satisfaction, dimension which was not significant different. The results for Amazon Alexa indicate that prior experience is not a strong determinant of VA UX. The lack of significant differences suggests that Alexa offers a consistent UX, irrespective of user prior experience with VA. One potential explanation is that Alexa offer a more user-friendly design and functionality, which render the advantages of prior experience less effective. Moreover, it is also noteworthy to consider the learning curve for Alexa is moderately flat, making it accessible to both experienced and non-experienced users. Alternatively, it is possible that Alexa's design prioritizes simplicity, ensuring a similar UX regardless of prior exposure. These results challenge the assumption that familiarity necessarily equates to an advantage in the context of Amazon Alexa.

In contrast to Amazon Alexa, the results for Google Nest VA demonstrate a more complex relationship between prior experience and UX dimensions. While experienced users scored significantly higher in the "Performance" and "Ease of Use" dimensions, the "Satisfaction" dimension did not exhibit a significant difference. This shows regardless of the prior user experienced, VA are found satisfying overall. Moreover, it also echoes a multifaceted relationship between prior experience, performance, and overall satisfaction. Hence, the results raise questions why experienced users got higher performance scores. Is it because they've gotten comfortable at using voice commands, or is it because they are familiar with the system capabilities? Figuring out why experienced users perform better can offer valuable insights into how to enhance UX for all users. The "Sentiment" dimension, which reflects users' emotional responses, revealed a noteworthy difference between experienced and non-experienced users for Google Nest, with experienced users expressing more positive sentiments. This findings shows experienced might influence not only the functional aspects of VA but also the emotional aspects. Additionally, User Expectations plays a role in UX, User expectations appear to be related to sentiment dimension. Since experienced users expressed more positive sentiments, this could be attributed to their prior experience aligning with their expectations, resulting in a more positive experience.

Now to delve deeper into our second hypothesis;

H2: Users with prior experience will have higher approval levels with the voice assistant's human-like conversational mannerisms compared to users with no prior experience.

The "Conversational Mannerism" dimension in the case of Amazon Alexa stands out as intriguing. Although the Levene's test hinted at unequal variances, the t-test did not reach statistical significance. The confidence interval, however, suggested a potential subtlety in user perceptions, hinting at the possibility of differences that are not readily apparent. This dimension warrants further investigation. The absence of statistical significance could indicate that users, regardless of prior experience, generally perceive Amazon Alexa's conversational mannerisms similarly. Nonetheless, when considering Google Nest VA, there is a significant difference between users with prior experience and those without. The findings suggest that experienced users found conversation mannerism more favorable supporting our H2 hypothesis, hinting that non-experienced users may not have been as acclimated to Google Nest voice interface. This observation contrasts how the users group measure Amazon Alexa VA conversational mannerism which reiterate the acclaimed voice user interface of Alexa.

This study also acknowledged its limitations, such as the small sample size, which could affect statistical significance of the mean difference. Additionally, the inherent variability in user interactions with VA and the evolving nature of these devices could impact results. Additionally, it's important to note that our research did not include tasks related to the Internet of Things (IoT), even though interacting with IoT devices is a common use of VA This limitation in our study may affect the broader applicability of our findings towards all VA tasks. Another limitation is our study focused on only two types of VA, and each of them produced different results. This emphasizes the type of VA might affect the UX, indicating that different VA may yield different outcomes. To make our findings more applicable across various VA technologies, broader range of VA should be included in future studies.

6 CONCLUSION

In conclusion, this study provides valuable insights into UX with Amazon Alexa and Google Nest VA. It highlights the influence of prior experience on user preferences and underlines the significant difference depending on the VA. Users, have distinct preferences for VA based on their individual prior experience and expectations. As VA continue to evolve and become integrated into various aspects of our lives, understanding how previous experience affect user experience becomes paramount. Manufacturers and developers must consider these nuances during design, in order to develop VA that cater to a wide range of users, from novices to experts. In future research, qualitative methods could be used to explore aspects such as how people interact with VA and the emotions they feel during these interactions. This deeper understanding can help us create a better-design for VA technologies and develop more effective marketing strategies. Additionally, it's essential to include other VA like Apple's Siri and Microsoft's Cortana in future studies. This will help us generalize our findings and also carry out a task by task analysis to understand how these assistants perform differently in various tasks. We should conduct a detailed analysis of these differences to gain a comprehensive understanding.

ACKNOWLEDGMENTS

The first author was supported by the Petchra Pra Jom Klao Ph.D. Research Scholarship from King Mongkut's University of Technology Thonburi.

REFERENCES

[1] Beirl, D. *et al.* 2019. Using Voice Assistant Skills In Family Life. Computer-Supported Collaborative Learning Conference, CSCL.ollaborative Learning Conference, CSCL. (2019), 96–103.
[2] Brüggemeier, B. *et al.* 2020. User Experience of Alexa when controlling music: Comparison of face and construct validity of four questionnaires. ACM International Conference Proceeding Series (Jul. 2020).
[3] Chae, J.-G. *et al.* 2019. Voice Assistant for Visually Impaired People. The Journal of Korean Institute of Information Technology. 17, 4 (2019), 131–136. DOI:https://doi.org/10.14801/jkiit.2019.17.4.131.
[4] Claessen, V. *et al.* Virtual Assistants. March 2017, 116–130.
[5] Dutsinma, F.L.I. *et al.* 2022. A Systematic Review of Voice Assistant Usability: An ISO 9241-11 Approach. SN computer science. 3, 4 (Jul. 2022), 267. DOI:https://doi.org/10.1007/S42979-022-01172-3.
[6] Dutsinma, F.L.I. *et al.* 2022. Personality is to a Conversational Agent What Perfume is to a Flower. IEEE Consumer Electronics Magazine. 12, (2022), 20–26. DOI:https://doi.org/10.1109/MCE.2022.3180183.
[7] Ehrenbrink, P. *et al.* 2017. Google now is for the extraverted, cortana for the introverted: Investigating the influence of personality on IPA preference. ACM International Conference Proceeding Series. December (2017), 257–265. DOI:https://doi.org/10.1145/3152771.3152799.
[8] Guerino, G. *et al.* 2021. Assessing a Technology for Usability and User Experience Evaluation of Conversational Systems: An Exploratory Study. (Apr. 2021), 463–473.
[9] Hong, G. *et al.* 2021. Voice assistants and cancer screening: A comparison of alexa, siri, Google assistant, and cortana. Annals of Family Medicine. 19, 5 (2021), 447–449. DOI:https://doi.org/10.1370/AFM.2713.
[10] Hornbæk, K. and Hertzum, M. 2017. Technology acceptance and user experience: A review of the experiential component in HCI. ACM Transactions on Computer-Human Interaction. 24, 5 (2017). DOI:https://doi.org/10.1145/3127358.
[11] Hoy, M.B. 2018. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. https://doi.org/10.1080/02763869.2018.1404391. 37, 1 (Jan. 2018), 81–88. DOI:https://doi.org/10.1080/02763869.2018.1404391.
[12] Huang, Y. 2023. Research on the Development of Voice Assistants in the Era of Artificial Intelligence. SHS Web of Conferences. 155, (2023), 03019. DOI:https://doi.org/10.1051/shsconf/202315503019.
[13] Iancu, I. and Iancu, B. 2023. Interacting with chatbots later in life: A technology acceptance perspective in COVID-19 pandemic situation. January (2023). DOI:https://doi.org/10.3389/fpsyg.2022.1111003.
[14] Jeong, Y. *et al.* 2019. Exploring effects of conversational fillers on user perception of conversational agents. Conference on Human Factors in Computing Systems - Proceedings. (May 2019). DOI:https://doi.org/10.1145/3290607.3312913.
[15] Karapanos, E. *et al.* 2009. User experience over time: An initial framework User Experience Over Time: An Initial Framework. May 2014 (2009). DOI:https://doi.org/10.1145/1518701.1518814.
[16] Lewis, J.R. 1992. Psychometric evaluation of the post-study system usability questionnaire: the PSSUQ. Proceedings of the Human Factors Society. 2, January 1992 (1992), 1259–1263. DOI:https://doi.org/10.1177/154193129203601617.
[17] Lewis, J.R. and Hardzinski, M.L. 2015. Investigating the psychometric properties of the Speech User Interface Service Quality questionnaire. International Journal of Speech Technology. 18, 3 (Sep. 2015), 479–487. DOI:https://doi.org/10.1007/s10772-015-9289-1.
[18] Li, S. *et al.* 2021. The PUEVA Inventory: A Toolkit to Evaluate the Personality, Usability and Enjoyability of Voice Agents. arXiv preprint. arXiv:2112.10811 (Dec. 2021).
[19] Liu, Q. Beyond technology use: A people-centred approach to reconceptualising the adoption of learning technologies. 3, 2022, 1–21.
[20] Lopatovska, I. 2020. Personality Dimensions of Intelligent Personal Assistants. Proceedings of the 2020 conference on human information interaction and retrieval (2020), 333–337.
[21] Maita, C.C. 2018. An Exploratory Study on Consumer Perceptions of Amazon Echo, Alexa, and Smart Speakers. Appalachian State University.
[22] Minge, M. *et al.* 2014. Exploring the Potential of Gameful Interaction Design of ICT for the Elderly. 2012 (2014), 304–309.
[23] Oh, Y.H. *et al.* 2020. Differences in Interactions with a Conversational Agent. International Journal of Environmental Research and Public Health 2020, Vol. 17, Page 3189. 17, 9 (May 2020), 3189. DOI:https://doi.org/10.3390/IJERPH17093189.
[24] Psychology, W. and Psychology, W. 2023. Transportation Research Part F: Psychology and Behaviour Acceptance of self-driving cars among the university

community: Effects of gender , previous experience , technology adoption propensity , and attitudes toward autonomous vehicles. 94, February 2022 (2023), 353–361. DOI:https://doi.org/10.1016/j.trf.2023.03.005.

[25] Reis, A. *et al.* 2017. Using Intelligent Personal Assistants to Strengthen the Elderlies' Social Bonds. (2017), 593–602. DOI:https://doi.org/10.1007/978-3-319-58700-4_48.

[26] Rohan, R. *et al.* 2022. Applying the Stimulus Organism Response Framework to Explain Student ' s Academic Self-concept in Online Learning During the COVID-19 Pandemic Applying the Stimulus Organism Response Framework to Explain Student ' s Academic Self-concept in Online Learnin. Springer Nature Singapore.

[27] Rohan, R. *et al.* 2023. Hey Alexa . . . Examining Factors Influencing the Educational Use of AI-Enabled Voice Assistants During the COVID-19 Pandemic. (2023), 23–28.

[28] Roshini, M. *et al.* 2022. AI System With Voice Modulation Using Neural. Ijraset Journal For Research in Applied Science and Engineering Technology. May (2022). DOI: https://doi.org/10.22214/ijraset.2022.42193.

[29] Studies, F. 2015. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. February 2000 (2015). DOI:https://doi.org/10.1287/mnsc.46.2.186.11926.

[30] Taylor, S. and Todd, P. 2013. Assessing IT Usage: The Role of Prior Experience. MIS Quarterly: Management Information Systems. 19, 4 (2013), 561–570. DOI: https://doi.org/10.2307/249633.

[31] Tiina Männistö-Funk, T.S. 2018. Voices from the Uncanny Valley. 4, 1 (2018). DOI:https://doi.org/10.14361/dcs-2018-0105.

[32] Tubin, C. *et al.* 2021. User experience with conversational agent: a systematic review of assessment methods. Behaviour and Information Technology. (2021). DOI:https://doi.org/10.1080/0144929X.2021.2001047/FORMAT/EPUB.

[33] Weinberg, B.A. and Weinberg, B.A. 2004. Experience and Technology Adoption. IZA DP No., 1051 (2004).

[34] Zhong, R. *et al.* 2022. User acceptance of smart home voice assistant: a comparison among younger, middle-aged, and older adults. Universal Access in the Information Society. 0123456789 (2022). DOI:https://doi.org/10.1007/s10209-022-00936-1.