"Alexa, are you spying on me?": Exploring the Effect of User Experience on the Security and Privacy of Smart Speaker Users

George Chalhoub[0000−0003−2082−2610] and Ivan Flechais

Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK
{george.chalhoub,ivan.flechais}@cs.ox.ac.uk

Abstract. Smart speakers are useful and convenient, but they are associated with numerous security and privacy threats. We conducted thirteen interviews with users of smart speakers to explore the effect of user experience (UX) factors on security and privacy. We analyzed the data using Grounded Theory and validated our results with a qualitative meta-synthesis. We found that smart speaker users lack privacy concerns towards smart speakers, which prompts them to trade their privacy for convenience. However, various trigger points such as negative experiences evoke security and privacy needs. When such needs emerge, existing security and privacy features were not found to be user-friendly which resulted in compensatory behavior. We used our results to propose a conceptual model demonstrating UX's effect on risk, perceptions and balancing behavior. Finally, we concluded our study by recommending user-friendly security and privacy features for smart speakers.

Keywords: User Experience · Smart Speaker · Security · Privacy · Behaviors.

1 Introduction

The first practical keyboard was invented by Christopher Latham Sholes in 1873 [21]. The keyboard and other peripheral devices were invented because traditional computing devices were not able to decode human voices. However, the rapid development of speech recognition technology is changing the way people interact with technology. Mobile phones are equipped with speech-activated functions supported with advanced and accurate speech-to-text technology. One of the most significant and successful voice technology innovations are smart speakers. Smart speakers like Amazon Echo, Google Home, and Apple HomePod are increasingly becoming a trend in homes and rapidly becoming integrated with other smart devices. Amazon's devices team announced in January 2019 that the company had sold more than 100 million Alexa powered devices worldwide [12]. In 2018, Google revealed that they sold more than one Google Home product every second. [19].

Smart speakers offer hands-free and eye-free operations allowing users to send voice commands while working on other tasks. Smart assistants like Alexa also

emulate social presence due to being equipped with speech synthesis technologies allowing Alexa to artificially produce human-like speeches [13]. To be able to operate in a hands-free environment, smart speakers need to continuously listen to what is being said around the device to catch the wake word (e.g., "Ok Google"). Cybersecurity critics have argued that always-on devices like smart speakers bring a significant threat to privacy and security. Smart speakers were previously vulnerable to security attacks which allowed attackers to turn them into a wiretapping device [13]. Integrating proper privacy and security controls into smart speakers while preserving UX seems to be a continuous challenge.

Users don't just look for privacy and security from those devices; they look for satisfaction, convenience, and well-being. They want to use technology with-out worry while having a good User Experience (UX). The UX of smart speakers involves much more than usability; it includes people's feelings, emotional reactions and psychological needs. The purpose of this research is to allow us to understand better how UX influences users' security and privacy. Therefore, we asked the research question: How do UX factors influence the security and privacy of users of smart speakers?

To tackle our research question, we conducted semi-structured interviews with thirteen users of smart speakers and analyzed the data with Grounded Theory. We summarize our key findings below:

- Users express a lack of privacy concerns towards smart speakers be-cause of individual perceptions (e.g., their perceived notability).
- Users trade their security and privacy for the benefits arising from smart speakers (e.g., convenience and utility).
- Users have various security and privacy needs that result from specific trigger points (e.g., detrimental experiences, adversarial needs).
- Common security and privacy features (e.g., muting) of smart speakers were not found to be user-friendly and were hindering the UX.
- Users reported compensatory behavior (e.g., disconnecting the devices, deleting audio history) resulting from negative experiences with smart security and privacy tools.

We used our results to present recommendations for the security and privacy design of smart speakers. In addition, we proposed a conceptual framework showing how UX interacts with risk and balancing behavior.

This research paper is organized as follows. Section 2 provides background in-formation related to UX and smart speakers. The section also discusses related works. Section 3 describes our study methodology and design. Section 4 presents a detailed description of our results, organized according to the discovered categories. Section 5 introduces design recommendations for security and privacy in smart speakers. In addition, it introduces our proposed conceptual framework. Finally, Section 6 presents our conclusion.

2 Theoretical Background

2.1 User Experience

Definition There is no universally accepted definition of UX. However, we will follow the definition by the international standard of human-system interaction ISO 9241-210, which defines UX as *"a person's perceptions and responses that result from the use or anticipated use of a product, system or service"* [23]. The definition includes a person's emotions, psychological responses, beliefs, perceptions, behaviors, preferences, and accomplishments.

Research Approach While the Human-Computer Interaction (HCI) community cannot agree on a uniformly accepted UX model that drives research, there is an agreement that UX is subjective, dynamic, and context-dependent [38]. UX research is mainly divided into two research methods: one method which advocates a qualitative design approach and one method that promotes a quantitative model approach. There are two prominent UX frameworks for each approach: McCarthy and Wright's approach [39] which is considered as qualitative design-based and Hassenzahl's approach [30] which is considered as quantitative model-based.

McCarthy and Wright's Framework McCarthy and Wright's framework [39] draws attention to the significance of a holistic experience view without reductionism [49]. The experience is described as holistic, dynamic, and subjective. The framework suggests threads that help describe experience based on context, time, feelings, emotions and processes which describe how a user subjectively makes sense of an experience.

Hassenzahl's Framework Hassenzahl's framework [30] focuses on the technological artifacts that affect the experience. The framework specifies distinct properties of experience (e.g. subjective, dynamic, holistic, situated). Based on the self-regulation theory by Carver and Scheier [16], the framework is composed of a tiered hierarchical UX model that describes experiences as being related to motives, actions, and specific conditions.

Factors of UX UX is influenced by three factors: user, system, and context [45]. These factors act as primary dimensions of UX, where sub-factors emerge from the literature. For the user factor, the related sub-factors that appear are emotions and psychological needs. As for the system factor, the sub-factors are hedonic and pragmatic product quality. For the context dimension, time and situatedness are the sub-factors.

Use in this paper For this paper, which is concerned with the security and privacy of smart home speakers, we will apply Hassenzahl's UX framework because it is concerned with the design of technological products and the UX.

2.2 Smart Speakers

Description of Smart Speakers A smart speaker is a wireless voice command device with a virtual assistant offering multiple hands-free services with the help of activation words known as wake-up words or hot words. Smart speakers consist of one or more microphones which await the wake-up word followed by a command from the user. Smart Speakers provide extra capabilities by allowing third-party developers to create applications that offer services. They can run If This Then That (IFTTT) automation applications that connects cloud services and users' devices. Smart speakers are associated with numerous security and privacy concerns [28, 33, 50], we summarize them below.

Security of Smart Speakers

Voice Authentication VAs allow users to communicate remotely by saying the wake-up word followed by the voice command. VAs had struggled in the past to recognize human voices, which prompted any audio within microphone range to send requests to the smart speaker. Smart speaker devices have numerous cases of interaction with television programs and advertisements. In 2017, a cartoon which included repeated Amazon Echo and Google Home commands had wrecked some of the viewer's devices [2]. Moreover, a notable attack on VAs is known as DolphinAttack [52] which sends voice commands in the form of an ultrasonic sound, a high-frequency sound that the human ear cannot detect.

Wiretapping Smart speakers are at risk of getting turned into wiretapping devices. Security researchers from Tencent demonstrated a security vulnerability at DefCon that would allow attackers to take complete control of the device, which would enable them to eavesdrop on private conversations [32]. Other related security vulnerabilities were discovered by Checkmarx [1] and MWR [10].

Voice Commands Smart speakers voice commands are transferred and stored in cloud servers [50]. While the data sent to the cloud is encrypted, it does not prevent a network sniffer from knowing there is an interaction happening with smart speakers [8]. Major smart speaker brands like Google Home, Amazon Echo, and Apple HomePod store the audio recordings in the cloud [41]. Google [36] and Amazon [25] allow users to manage and listen to their audio activity online, which adds a security risk in case an account is compromised [50].

Privacy of Smart Speakers

Company Monitoring Major smart speaker companies (e.g., Google, Apple and Amazon) employ staff to manually listen to consumers' voice commands to improve speech recognition technology [17]. A Bloomberg investigation revealed that Amazon had contracted thousands of humans to work in a secret program with each employee processing up to 1,000 audio clips in 9-hour shifts [22]. Amazon responded by saying that their contractors do not have access to customers'

personally identifiable information [17]. However, critics argued that contractors may have access to GPS coordinates which can be used to point to users' locations [46]. In addition, companies collect personal information such as names, IP addresses, locations, addresses and payment cards [3]. German magazine Heise reported the story of an Amazon customer who decided to exercise his GDPR rights by requesting his stored personal information from Amazon. The company mistakenly sent 1,700 audio files and a transcribed document containing the interactions of users with Amazon Echo [11].

Misinterpretation Misinterpretation of wake words and commands raise privacy concerns. An investigation by Symantec revealed that wake-up words could trigger smart speakers even if they are not accurate. The research reports that Google Home woke up for 'Ok Bobo' instead of 'Ok Google' [50]. A case reported by KIRO7 confirms the finding where a family in Portland had their private conversation recorded by Alexa and sent to a random contact due to misinterpretation [31].

Law Enforcement Smart speakers collect and store a massive amount of personal information, prompting law enforcement to often demand access to data. A double murder investigation in New Hampshire prompted a judge to order Amazon to submit any audio recordings by Echo during the day of the murder [27]. Prosecutors have also sought evidence from Amazon Echo in a case involving the killing of an Arkansas police officer [14]. To protect consumer privacy, Amazon filed a motion against the police warrant issued by the prosecutors [15] but later released the data once the owner of the Echo consented [40]. Although Amazon was able to fight off the judge's orders, some privacy experts warn that laws can be passed to allow law enforcement to remotely activate smart speakers and eavesdrop on suspects [20].

2.3 UX of Smart Speakers

Unlike laptops and mobile phones, smart speakers do not generally have a screen. Even with screen-enabled smart speakers, the interactions remain invisible and the designers often aim at a positive 'voice experience'. The lack of visuals used in interactions makes designing and measuring UX more challenging [5]. Pyae and Joelsson conducted a web-based survey with 114 users and found that Google Home devices result in positive UX but had some usability issues [44]. There are current issues with understanding UX design for Voice Assistants (VA) of Smart Speakers [34]. Traditionally, measuring user satisfaction consisted of analyzing clicks and scroll signals. However, those signals do not exist in smart speakers which makes it challenging to measure user satisfaction. Other researchers have proposed ways to measure new signals. For instance, Hashemi et al. [29] proposed user intent as an original signal for measuring user satisfaction. Moreover, the personification of Alexa is linked to a higher level of user satisfaction due to increased social interactions [43]. The personification of VAs might require UX designers to work with Machine Learning as a design material [34].

2.4 Related Work

Lau et al. [37] ran a diary study and semi-structured interviews with 17 users and 17 non-users of smart speakers to understand users' reasoning for the adoption of those devices, privacy concerns and insights, and experiences. Smart speaker users were found to have a sophisticated trust relationship with companies behind smart speakers, a lack of complete understanding of privacy risks and a dependence on the socio-technical context where smart speakers are. The researchers also found that users rarely use the privacy features of smart speakers. Moreover, non-users expressed distrust for smart speakers' companies and did not find smart speakers useful. Pascal Kowalczuk [35] analyzed more than 2,000 customer reviews and 850 tweets and found that enjoyment has the largest effect on the intention to use smart speakers. Other factors that strongly adopted the use of smart speakers were found to be: usability, equality and diversity of the product, consumer's technology optimism, and the security and privacy risk. Yang et al. [42] ran a questionnaire for 315 individuals in South Korea to study user intentions for adopting smart speakers. They found that the risk of smart speaker use did not have a significant effect on the perceived value of speakers. The authors tried to justify the findings with two possible explanations. The first explanation is that privacy is the major viewed risk in speaker adoption which could have a negligible effect on the perceived value [51]. The second explanation is that smart speaker users may not be knowledgeable of all the risks associated with smart speakers. To the best of our knowledge, there has been no previous work that investigates the role of UX in security and privacy in smart speakers.

3 Research Methodology

Our study aims to explore how UX factors affect security and privacy in smart speakers; therefore, we used a qualitative research approach (Figure 1). Our approach consisted of collecting data using semi-structured interviews. This exploratory approach allowed us to reveal new information from participants and uncover UX factors (e.g., emotions and motivations).

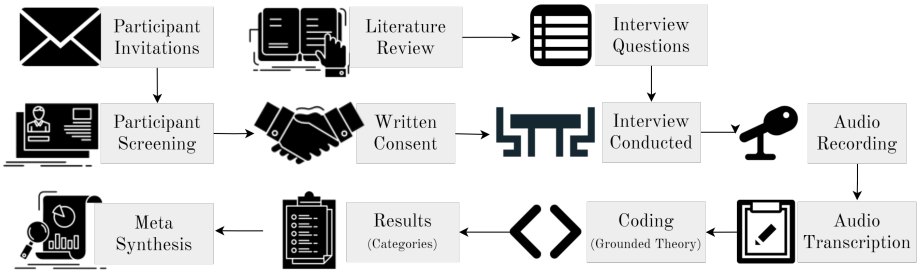

Fig. 1. Summary of our research methodology

3.1 Recruitment

To recruit participants, we printed recruitment flyers and posted them in different department buildings. We also published announcements in local city forum posts (e.g., our city's local subreddit [18]). Furthermore, we sent recruitment emails for participants using university-provided mailing lists. The recruitment message contained eligibility criteria and contact details. Initial communication with potential participants happened via university email.

3.2 Sampling

We used purposive and theoretical sampling to recruit a sample of thirteen smart speaker users to participate in our research study. Purposive sampling allowed us to select specific eligible participants from preselected criteria. The eligibility criteria consisted of users who: (i) were at least 18 years, (ii) used smart speakers in the past three months, (iii) were able to communicate in English and (iv) were able to give consent. Theoretical sampling allowed us to inform the sample size (n=13) which was determined based on theoretical saturation. We performed data analysis after each interview and we stopped recruitment when interviews did not provide any additional categories. The demography of the participants is summarized below (Table 1).

Table 1. Participant Demographics

ID	Age Group	Education	Gender	Device
P1	25-30	High School	Female	Google Home
P2	30-35	High School	Male	Amazon Echo Dot
P3	35-40	Bachelors	Male	Amazon Echo Dot
P4	20-25	Bachelors	Male	Google Home Mini
P5	20-25	Doctorate	Male	Google Home
P6	20-25	Masters	Male	Google Home, Apple HomePod
P7	35-40	Bachelors	Male	Amazon Echo Dot
P8	20-25	Masters	Male	Google Home Mini
P9	25-30	Masters	Male	Amazon Echo Dot
P10	40-45	Masters	Female	Amazon Echo, Amazon Echo Dot
P11	20-25	Bachelors	Female	Amazon Echo Dot
P12	25-30	Masters	Male	Amazon Echo
P13	25-30	Bachelors	Male	Amazon Echo

3.3 Data Collection

Interviewees were invited to attend the interview in person. The interviews were conducted within interview rooms in university buildings. Four participants could not be present and were interviewed via Skype. The interview questions were based on the literature review conducted and tackled topics related to UX

factors. All the interviews were audio-recorded using a recording device. Written notes were taken during the interview. The length of the interviews varied between 28 minutes and 62 minutes. All the participants were thanked with a £10 ($12) Amazon gift card voucher regardless of whether they completed the interview or not.

Interview Process The experimenter first started with collecting necessary information from interviewees such as their age, gender, education, employment. Interviewees were then asked about the number and type of smart speakers that they use. The experimenter had deeper probing about the environment of the smart speaker. Interviewees were asked to justify all of the decisions they have made such as reasons for using a smart speaker, picking a particular brand and placement of the speaker in a particular location. Interviewees were then asked to explain how they understand the technology behind smart speakers and discuss any unpleasant interactions. This was followed by an open-ended discussion of situations where the interviewees felt uncomfortable or uneasy around the smart speaker. Based on the previous experiences and knowledge of interviewees, circumstances related to privacy and security were further explored.

3.4 Data Analysis

All the recorded interviews were transcribed and repeatedly read for familiarization with the present data. We used Grounded Theory to analyze our data. Interviews were coded with data analysis software Nvivo 12.0. At the end of the analysis, we identified 127 codes. To validate our findings, we consolidated the existing literature and used meta-synthesis [48] to compare our results with the reviewed literature.

3.5 Limitations

We have interviewed smart speaker participants who clearly chose to use and adopt smart speakers. Users of smart speakers are not a representative of all users. Non-users are likely to have different views and perceptions.

3.6 Ethics

Oxford University's Central University Research Ethics Committee reviewed and approved our research study (CUREC/CS_C1A_19_024). At the beginning of each interview session, we gave each participant an information sheet and a consent form which they had to sign before taking part in our study.

4 Results

We extracted six categories (Table 2) from our analysis (Figure 2).

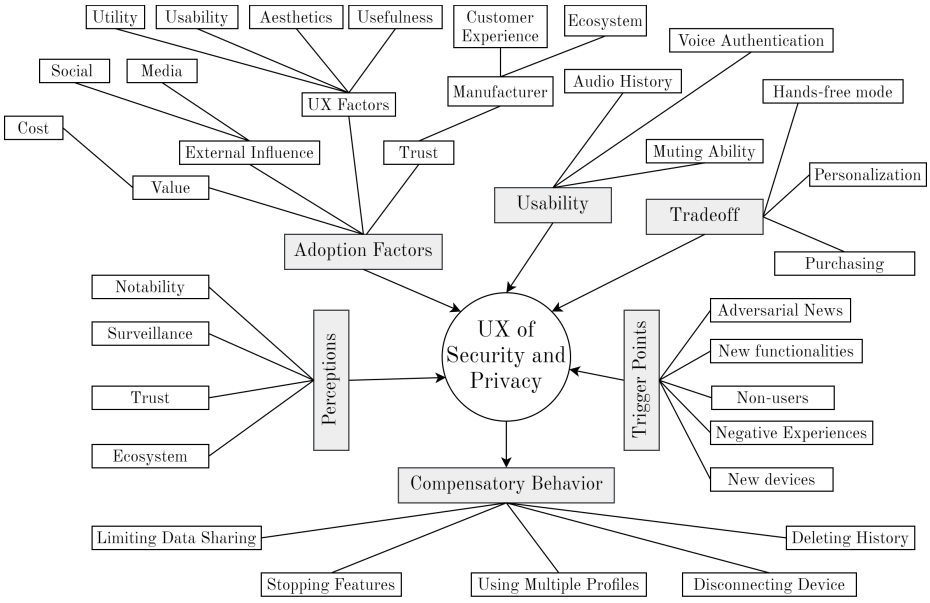

Fig. 2. Summary of our categories and codes

Table 2. Summary of extracted categories

Perceptions and beliefs towards privacy resignation
Perceptions leading to privacy resignation: perceived notability, government surveillance, trust, and product ecosystem.
Usability and pragmatic quality of security and privacy controls
Usability of smart speaker's security and privacy controls: muting ability, voice authentication, and audio recording history.
Influencers in the trade-off between privacy and convenience
Features affecting the trade-off choice between privacy and convenience: personalization, hands-free mode, and purchasing.
Factors and motivators affecting smart speaker adoption
Factors determining smart speaker adoption: usefulness, trust, hedonic quality, cost, and social influence.
Trigger points for security and privacy considerations
Occasions prompting security and privacy considerations: adversarial news, non-users and negative experiences.
Security and privacy compensatory behavior
Reported compensatory behavior: limited use, disconnecting the device, stopping audio history and using multiple profiles.

4.1 Perceptions and beliefs towards privacy resignation

Users express different perceptions and beliefs towards giving up their personal data to their smart speakers. We identified four perceptions and beliefs:

Perceived notability Users of smart speakers are influenced by how notable they think they are. When discussing giving personal data to the speaker, five users said they're not concerned about data collected by smart users because they have nothing to hide. Other users said they do not feel targeted by any external entities. When asked about concerns regarding their data being stolen, two participants responded by saying they are not an interesting target and don't feel targeted as a result. P5 said: "*I think it's easy to kind of get wrapped up in worrying about being followed or being tracked online. But in reality, probably not going to happen to us. We're not a person of particular importance.*"

Surveillance Some participants dismissed privacy concerns since they believe that government and corporate surveillance can obtain their personal data. Quoting P7: "*At the end of the day, if government agencies want to see what I'm doing, they can. I'll never know. So, what's the point of worrying about it?*" Also, some participants dismissed smart speaker microphone concerns because they claimed they are no different than their smartphones. Quoting P6: "*Why does one smart speaker microphone make a difference? Some people wouldn't talk around Alexa because it seems like an over-listening device. But also, ultimately, it is not that different from smartphones.*"

Trust All thirteen participants said that they trust their smart speaker manufacturer (e.g., Google, Apple, Amazon) to secure their personal data. As a result, they feel safe using the devices despite some saying that the companies might use it for *"targeted advertising"* (P6) and *"commercial gains"* (P1).

Ecosystem Some participants dismiss privacy concerns because their data is shared with the smart speaker's manufacturer through their ecosystem. P6 who massively uses Google's services (e.g., Gmail, Drive, Photos) thought that adding Google Home won't make a difference. P6 said: "*I did think of the privacy of it. But once I saw how it was being used on, I thought about this whole Google ecosystem which I'm already tied into, I thought well*". Similarly, P6 had used Amazon services for more than two decades and was comfortable using the Echo Dot in their home. Quoting P6: "*Amazon must have an incredible profile on me because I've used it for the last 20 years, they have a total profile of what my hobbies are, what I like and what I don't like. So, I don't care. Really.*".

4.2 Usability and Pragmatic Quality of Security and Privacy Controls.

We explored the usability of common security and privacy controls.

Muting ability Some users wanted to mute the smart speakers for privacy reasons but were frustrated because the devices can only be physically muted. Quoting P10: "*This is unhelpful. Echo devices are on high shelves. I can't just reach up and click it. I have to actually go and get it and pull it down and then press it. Being able to voice control would be more useful*". Other participants went further by suggesting that they would be annoyed if the smart speaker is remotely muted because they will need "*to get up to unmute it, because it is not listening anymore*" (P11). P1 said they would prefer to have a temporary remote mute feature that would mute the device for a short period: "*I wish there was a feature where you tell Google not to listen to you for like 10 minutes and it starts listening to you again after 10 minutes.*"

Audio Recording History Most Amazon Echo users know that they can view their audio recordings using the Amazon app. Two participants said that they regularly delete their audio recording history as part of digital hygiene or housekeeping. Three participants described their stored history as "*pointless*". Two participants who use the Google Home said that they wanted to check their queries online; however, they found the process to be complicated and confusing. Quoting P4 "*You needed to do like 7-8 steps to be able to see your voice commands. After a few minutes, I gave up.*".

Voice Authentication Echo users expressed feelings of trust and security towards ordering from Amazon due to the Echo's Purchase by Voice feature. The feature prompts Alexa to individually recognize voices using 'Alexa Voice Profiles' and reportedly is easy to set up and effortless. Quoting P3: "*It was easy to set up, Alexa made me say a couple of things and then it easily worked. If someone tries to use the Alexa in my house to order things, they won't be able to, because the voice thing will be able to block it.*" Google Home's voice authentication feature was not supported for UK households during the time of the interview.

4.3 Trigger points for security and privacy considerations

We identified trigger points prompting users to re-consider their security and privacy.

Adversarial news Adversarial news originating from news stories or social contacts tend to prompt smart speaker users to consider what they share with the device. User P9 recalled a news article about Amazon: "*You could read in the past that Amazon had some issues with the data, for example, gave data from one person A to person B. They didn't even know each other*". In addition, P9 felt worried after finding a news article alleging that Alexa would recognize if they were ill.

New functionalities New smart speaker functionalities might prompt users to question whether they would use smart speakers. While one participant had the Echo Show 5, which contains a camera, most participants were not comfortable with using a smart speaker with a camera. P11 considers microphones to be less concerning: "*It's just cameras. It's like having CCTV in your home. You don't want people watching you eat peanut butter at 3 am in the morning. It's a bit more concerning, I guess. Audio is less concerning than video for sure*". In addition, when asking participants whether they would bank with their device, many have completely dismissed the idea.

Non-user Non-users of smart speakers prompt some users to consider their privacy around the device. P1 warns his guests about the device: "*I would tell my guests that the Google Home is listening to them. You know, if they have anything very private to say, or if they would want me to mute it, then I would mute it.*" P2, who possessed multiple Echo Dots at home and work, started having considerations about leaving it active when co-workers are around. P2 said that they have never muted the device at home but when they began using it at work, they thought that it is appropriate to mute it. Similarly, P13 expressed similar behavior when they had their client visiting them at home.

Negative Experiences Some users reported negative experiences during their use of smart speakers, which prompts them to consider their behavior. Participant P8 who had difficulties checking his Google Home audio log was able to review his logs eventually and discovered that multiple non-intended conversations were recorded. Quoting P8: "*I really thought the Google Home was innocent and all. Until I realized that a lot of unintended conversations were recorded, yikes*". Another negative experience reported by P10 relates to the use of the purchasing feature by Amazon Echo. P10 discovered later that their son had made multiple orders from Amazon by tweaking the device settings. P10's negative experience prompted them to consider whether the purchasing feature on their device is secure enough and whether it should remain activated.

Acquiring New Devices Acquiring a new smart speaker for the first time might be a trigger point for privacy considerations. Participant P4 explained how receiving a Google Home as a gift triggered a privacy consideration: "*I didn't want to get a smart speaker. And when I got it as a gift, I just kept it in the drawer. Then I thought: Hey, it's not recording me randomly. Why would it be? And then, one time, I just put it on and slowly got over the fear of using them*".

4.4 Factors and motivators for smart speaker adoption

We discovered six major factors and motivators for smart speaker adoption:

Usefulness Usefulness is the most common factor for smart speaker adoption. Before acquiring smart speakers, ten participants anticipated that the device will be useful, convenient, and will "*make life easier*" (P13). P1 purchased Google Home to be able to ask the assistant for quick questions: "*I thought the Google Home would be well equipped to answer my queries quickly.*" Other widespread purposes that users anticipated to be very useful were: playing music, managing their calendar, checking the weather, messaging and getting the news.

Trust Participants' trust for smart speaker manufacturers affects whether they would adopt a smart speaker or not. P2 would not have purchased a smart speaker if Google was the only company that manufactured those devices because they don't trust the company. P2 said, "*I really trust Amazon as a company, I've used many of their services before*". In contrast, P5 trusts Google said "*I like Amazon a lot actually, in terms of products and services. But I don't trust them as much as I trust Google*".

Aesthetic and Hedonic Quality The perceived aesthetic and hedonic quality of smart speakers influences their adoption. Before purchasing the product, P13 watched online videos and felt that the '*humanized voice of Alexa*' is satisfying. Not only were the aesthetics considered, but the size, looks and feels. Another user said that they were positively surprised by how small the Echo Dot and they thought the small device can easily hide out of sight if needed. Other reported qualities that were considered are the audio quality of the device, as well as the color and mobility.

Cost The cost of smart speakers seems to play a significant factor in acquiring and adopting smart speakers. Eight participants had either got smart speakers for free or paid a small amount during a sale period. Participant P10 "*won one*" while P4 "*got it as a gift*". Other participants acquired the device during sales such as "*black friday sales*" (P11), "*prime day*" (P13) or during a "*promotion*" (P8). Participant P3 was torn between getting Amazon's Echo Dot or Apple's HomePod, but after finding a promotion online for the Echo Dot, they made their decision: "*The Apple stuff is too expensive. We got a deal for the Echo Dots for 30 quid*". Two participants said they would not have purchased their smart speaker device at the usual price sold.

Social influence Social contacts who own smart speakers seem to influence non-users into acquiring them. P6 bought their own Google Home after a Google Home Mini was set up at their family's house. Similarly, P11 purchased their own device after they used the smart speaker of their partner a couple of times. P12 saw an Echo Dot at his cousin's residence before getting one: "*When I was at his place once, it looked like a very compact tool to have, I got jealous, and I thought that's a device that would like to have*".

Media Mass media also seems to influence or motivate users to purchase and use smart speakers. Two participants heard about smart speakers on the news before acquiring them. Quoting P7: "*I read an article in the newspaper and it said the next third generation of the Echo Dot is out. I saw something in the paper that was like, very interesting. I just thought this is going to be pretty cool. Actually, I was just kind of intrigued*". Similarly, participant P1 had watched videos and read about the Google Home before making the purchase.

4.5 Security and Privacy Compensatory Behavior

Users reported different cases of compensatory behavior.

Deleting Audio history When P8 went through his Google Home audio commands history and reviewed their audio history, the discovery of accidental recordings triggered a compensatory behavior. Unintended conversations could be recorded by the accidental triggering of the smart home assistant (e.g., mishearing the wake words). After this experience, P8 mentioned that they regularly review and monitor audio commands and delete queries that are considered to be non-intended or malicious.

Stopping device features P10's negative experience of having unauthorized purchases on their smart speaker from their child prompted a compensatory behavior. P10 had contacted Amazon customer service and was able to turn off the purchasing feature from their smart speaker: "*I was able to chat with customer support and completely stop this feature from working on my Alexa.*" In that case, P10 had a negative experience that caused them to lose money, and this has led them to take a course of action and stop this feature from their Alexa device.

Disconnecting the device Another reported example of compensatory behavior involved participant P13 and his client. They were having a regular discussion at P13's residence, which ought to be private and confidential. P13 had noticed that their client seemed very uncomfortable after spotting that the Google Home's LED Light showing "*running lights in white color*" which meant that Google Home is listening. P13 described the situation as very "*awkward.*" After facing this experience, P13 disconnects his smart speakers whenever they have a client visiting: "*We never discussed the matter. But whenever they are in my home, I make sure to plug off all the smart assistants*".

Using multiple Profiles Two participants set separate profiles for security or privacy reasons. P3 had enabled different profiles on their account to be the only person able to make purchases on the Alexa app. Quoting P3: "*So they're there, attached to me and set so that only I could make purchases through them.*" Another participant set up profiles on the Google Home to be able to

receive personalized results on that without feeling uncomfortable. Personalized results include data from Google apps such as Photos, Calendar, Contacts, and Purchases [26].

Limiting Data Sharing P4 described themselves as "*cautious*" when using their Google Home. In particular, when sending a command to the device, they make sure no compromising information is sent. Quoting P4: "*I make sure I don't say anything risky when it is recording. You know, I'm not going to, like, say my SSN out loud when it's talking.*" Some participants do not completely adopt smart speakers. They express reservations when looking at different features. For instance, P11 said they would never use the purchasing feature in the device, whereas P13 said they refuse to give the Alexa app access to their iPhone's list of contacts.

4.6 Deliberations in Privacy/Security and UX Trade-off

Personalization Smart assistants like Alexa, Siri, and Google Assistant are personalized; they tend to use customer's data and audio log to provide a personalized experience with the device. We asked users if they prefer a neutral smart assistant that does not store any of their personal data, which might reduce the UX with the smart speaker. Only one user said they wish to have a non-personalized assistant. Most participants said they prefer smart assistants that are personified, personalized and integrated into their daily lives. Some participants express numerous positive emotional reactions that heavily influence their trade-off choice. Quoting P12: "*I feel cognizant of the fact that sometimes I refer to the device as "the device". But sometimes I'll refer to the device as "she" or "her". Kind of like humanizing the device in a sense.*" Many users utilize their smart speakers daily for different tasks at different times of the day and the devices seem to be integrated into their lifestyle. P10 discussing personalization: "*Alexa almost feels like a member of the family and we just love her. We want her to stay smart and remembering our details*".

Hands-free mode. We asked participants if they prefer a version of smart speakers without the always-listening mode. 12 out of 13 participants dismissed the idea. When examining the trade-off between privacy and UX, they chose to sacrifice privacy for their comfort. Participants described a not-always-listening mode smart speaker as "*bothersome*" (P5), "*annoying*" (P1), "*a hassle*" (P12), "*difficult*" (P9) and "*defeating the purpose*" (P7) (P11). For disabled users, having a not always-listening mode could significantly impact their comfort. P10 weighted in "*I like the fact that I can wake up and ask what to do Alexa with my voice because I'm disabled, I can ask Alexa dozens of things to do for me without having to find my phone or another human being.*".

Purchasing Both Google and Amazon allow users to purchase items through smart speakers. Some non-users of this feature said that they don't trust the

whole process of buying via the device. One participant was okay with using for their smart speaker for small purchases but some felt "*uncomfortable sometimes for not knowing what is happening behind the scenes. Where is my credit card stored? What if they overcharged me?*" (P2). Participants who order via smart speakers expressed positive feelings, good UX, and trust towards purchasing and using the devices. Some users expressed feelings of trust and security from ordering off Amazon due to the Echo's 'Purchase by Voice' feature. The feature prompts Alexa to recognize voices using "Alexa Voice Profiles" and as a result, only allows the smart speaker owner to order from Amazon.

5 Discussion

5.1 Privacy Design Recommendations

Improvements to muting Amazon Echo and Google Home users cannot mute their smart speakers remotely (e.g., Alexa stop listening), which creates an inconvenience. For instance, disabled users suffer from significant disadvantages for not being able to mute their devices remotely. Manufacturers should add a device feature allowing users to remotely mute their speakers. Remote muting would require a physical trigger to unmute the device. Therefore, the feature should be accompanied by two complementary functions: *Temporary Remote Mute* and *Mobile App Unmute*. *Temporary Remote Mute* would allow participants to mute the speaker for a period of time (e.g., 'Ok Google, stop listening for the remainder of the day.'). *Mobile App Unmute* would allow users to unmute their devices via their mobile applications. Manufacturers of such applications should ensure that unmuting from apps is straightforward and easy to use (e.g., using GUI on/off toggle components).

Support for multiple devices It is not unlikely for household users to own multiple smart speakers. Having to remotely mute every device by voice may decrease the usefulness and usability. Manufacturers should support muting all (e.g., Hey Google mute all devices) or part of household devices from one device (e.g., Ok Alexa mute the living room speakers).

Changing Privacy Default Settings Google and Amazon store the audio history of their customers' commands by default. Google activates the 'Voice & Audio Activity' feature by default storing all of the customer's recordings. Similarly, Amazon turns on two features by default which permits their contractors to manually review a portion of the audio recordings. A significant number of our interviewees were not aware that their audio recordings are cataloged and stored. It seems highly unlikely that smart speaker users will go through the settings and disable features that pose a risk to their privacy or security (e.g., consenting to human review of their audio activity). Companies should ensure that privacy-preserving settings are switched on by default.

Improvements to the audio logs feature. While Google allows its customers to switch off the audio log activity feature, Amazon does not [24]. Users who do not want to have their audio activity stored would still need to delete their log from the device regularly – which would result in decreased UX.

Private Mode Some users mentioned that they would like to keep their audio recordings for practical reasons, which would increase the UX. Smart speaker manufacturers should introduce a private mode that is equivalent to the private mode of a web browser. Users who wish to have their activity logged could temporarily pause activity logging using the suggested private mode feature. The private mode could be complemented with two additional associated features: *Voice Activation* and *Associated Colors*. *Voice Activation* allows users to toggle the private mode by voice (e.g., Hey Alexa, turn on private mode). *Associated Colors* would change the color of the speaker to a specific color (e.g., red) when private mode is on.

5.2 Security Design Recommendations

Adding Security Layers to Voice Recognition Voice Recognition technologies have a history of security vulnerabilities (e.g., voice impersonation attacks [4]). Many of our interviewees had difficulties trusting the voice recognition features available on smart speakers – Google uses 'Voice Match' whereas Amazon uses 'Voice Recognition'. Smart speaker companies can add additional security layers to voice recognition (e.g., asking for memorable passphrases) – which is likely to increase the security and nurture trust.

Offline Capabilities While some participants use their devices for multiple and varied tasks, some report minimal use of the devices. Two participants have suggested they would like to use offline smart speakers. One of the participants' uses of their smart speaker is limited to controlling their smart home. The three major commercial smart speakers send every user query to the cloud for processing even if the command was straightforward (e.g., 'Alexa, shut off the lights').

Creating an offline smart speaker for performing basic tasks is possible. The company Sensory has developed an offline smart speaker that does not require any internet access. The device can perform voice recognition offline and perform many tasks such as setting the timer, control smart homes and playing music via Bluetooth [6]. Offline smart speakers nearly eliminate the security and privacy risks associated with cloud smart speakers.

5.3 UX Conceptual Model

Our results show that UX qualities (e.g., findable, desirable, credible) influence security and privacy in three areas: the perception of risk, the experience of harm and the mitigation practice. To present a model showing how UX affects behavior, we explored John Adam's theory of risk compensation, which states

that there is a "*risk thermostat*" influencing human behavior. The theory explains that users experiencing a safe lifestyle eventually seek out risky behavior; but overcompensate before returning to safety [7, 47]. Using the risk thermostat and our study findings, we proposed a conceptual model demonstrating how UX qualities interact with the concepts on risk and balancing behavior. In our model, the experience [39] of impact, vulnerability, and threat strongly influence users' perceptions of risk which would affect balancing behavior (Figure 3).

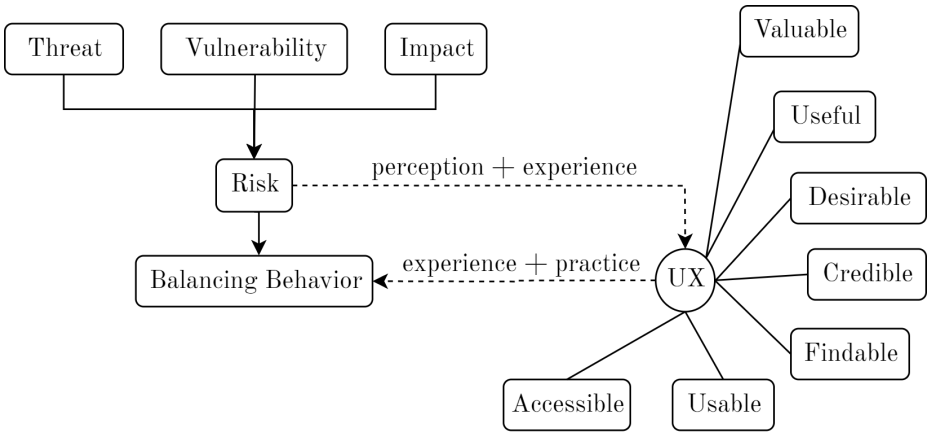

Fig. 3. Conceptual model demonstrating UX effect on risk and balancing behavior

6 Conclusion

With over a quarter of American adults owning a smart speaker [9], there is no doubt that smart speakers are witnessing considerable growth today. Smart speakers bring convenience and benefits to their users, but security and privacy concerns may be damaging their market growth. To find out how UX factors affect the security and privacy of smart speaker users, we conducted semi-structured interviews with thirteen users of smart speakers. We found that users reported compensatory behavior due to security and privacy features that were not user friendly. We used our results to recommend enhanced security and privacy features for smart speakers. Finally, we proposed a conceptual model that illustrates how UX qualities are linked with the concepts of risk and balancing behavior.

Acknowledgments The interview gift cards were funded by Oxford University's Centre for Doctoral Training in Cyber Security. The author of this research paper is funded by a grant from Fondation Sesam, a Geneva-based foundation.

References

1. Amazon Echo: Alexa Leveraged as a Silent Eavesdropper. Tech. rep., Checkmarx, https://info.checkmarx.com/wp-alexa
2. 'South Park' Episode Triggers Viewers' Amazon Alexa and Google Home, vol. 2019. https://www.hollywoodreporter.com/live-feed/south-park-premiere-messes-viewers-amazon-alexa-google-home-1039035
3. What data does Amazon collect and use? , vol. 2019
4. Security Vulnerabilities of Voice Recognition Technologies, vol. 2019 (2015), https://resources.infosecinstitute.com/security-vulnerabilities-of-voice-recognition-technologies/
5. "Alexa, How Will Voice Impact User Experience?" (2018), https://userbrain.net/blog/alexa-voice-user-experience
6. Sensory is Enabling Offline Smart Speakers with No Cloud Connectivity to Maximize Security, vol. 2019 (2019), https://voicebot.ai/2019/01/18/sensory-is-enabling-offline-smart-speakers-with-no-cloud-connectivity-to-maximize-security/
7. Adams, J.: Risk and morality: three framing devices. Risk and morality pp. 87–106 (2003)
8. Apthorpe, N., Reisman, D., Feamster, N.: A smart home is no castle: Privacy vulnerabilities of encrypted iot traffic. arXiv preprint arXiv:1705.06805 (2017)
9. Auxier, B.: 5 things to know about Americans and their smart speakers, https://www.pewresearch.org/fact-tank/2019/11/21/5-things-to-know-about-americans-and-their-smart-speakers/
10. Barnes, M.: Alexa, are you listening?, vol. 2019 (Aug 2017), https://labs.mwrinfosecurity.com/blog/alexa-are-you-listening
11. Bleich, H.: Alexa, Who Has Access to My Data?
12. Bohn, D.: Amazon says 100 million Alexa devices have been sold. The Verge (2019)
13. Burke, S.: Google admits its new smart speaker was eavesdropping on users. Online at https://money.cnn.com/2017/10/11/technology/google-home-mini-security-flaw/index.html.Citation on p. 7 (2017)
14. Carman, A.: Police want an Echo's data to prove a murder case, but how much does it really know?, vol. 2019 (2016), https://www.theverge.com/2016/12/27/14089836/amazon-echo-privacy-criminal-investigation-data
15. Carman, A.: Amazon says Alexa's speech is protected by the First Amendment, vol. 2019 (2017), https://www.theverge.com/2017/2/23/14714656/amazon-alexa-data-protection-court-free-speech
16. Carver, C.S., Scheier, M.F.: On the self-regulation of behavior. Cambridge University Press (2001)
17. Cellan-Jones, R.: Smart speaker recordings reviewed by humans (2019), https://www.bbc.com/news/technology-47893082
18. Chalhoub, G.: r/oxford - Do you own a smart speaker (Alexa, Google Home)? Get paid £10 for a 30 minutes interview, https://www.reddit.com/r/oxford/comments/cgrxgm
19. Chandra, R., Huffman, S.: How Google Home and the Google Assistant helped you get more done in 2017. blog (2018)
20. Cranz, A.: Amazon's Alexa Is Not Even Remotely Secure and I Really Don't Care, vol. 2019. https://gizmodo.com/alexa-is-not-even-remotely-secure-and-really-i-dont-car-1764761117

21. David, P.A.: Clio and the Economics of QWERTY. The American Economic Review **75**(2), 332–337 (1985)
22. Day, M., Turner, G., Drozdiak, N.: Amazon Workers Are Listening to What You Tell Alexa. Retrieved June **27**, 2019 (2019)
23. DIS, I.: 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems (formerly known as 13407). International Standardization Organization (ISO). Switzerland (2010)
24. Fussell, S.: Consumer Surveillance Enters Its Bargaining Phase, vol. 2019 (2019), https://www.theatlantic.com/technology/archive/2019/06/alexa-google-incognito-mode-not-real-privacy/590734/
25. Garun, N.: How to hear (and delete) every conversation your Amazon Alexa has recorded, vol. 2019 (2018), https://www.theverge.com/2018/5/28/17402154/amazon-echo-alexa-conversation-recording-history-listen-how-to
26. Google: Allow personal results on your shared devices . https://support.google.com/assistant/answer/7684543
27. Hamilton, I.A.: A judge has ordered Amazon to hand over recordings from an Echo to help solve a double murder case, vol. 2019. https://www.businessinsider.com/amazon-ordered-to-disclose-echo-alexa-recordings-murder-case-2018-11
28. Hart, L.: Smart speakers raise privacy and security concerns. Journal of Accountancy **225**(6), 70 (2018)
29. Hashemi, S.H., Williams, K., Kholy, A.E., Zitouni, I., Crook, P.A.: Measuring User Satisfaction on Smart Speaker Intelligent Assistants. Anne Dirkson, Suzan Verberne, Gerard van Oortmerssen & Wessel Kraaij p. 22 (2018)
30. Hassenzahl, M.: Experience design: Technology for all the right reasons. Synthesis lectures on human-centered informatics **3**(1), 1–95 (2010)
31. Horcher, G.: Woman says her Amazon device recorded private conversation, sent it out to random contact, vol. 2019 (2018), https://www.kiro7.com/news/local/woman-says-her-amazon-device-recorded-private-conversation-sent-it-out-to-random-contact/755507974
32. HuiYu, W., Wenxiang, Q.: Breaking Smart Speakers: We are Listening to You., vol. 2019. https://www.defcon.org/html/defcon-26/dc-26-speakers.htmlHuiYu
33. Jackson, C., Orebaugh, A.: A study of security and privacy issues associated with the Amazon Echo. International Journal of Internet of Things and Cyber-Assurance **1**(1), 91–100 (2018)
34. Kaye, J., Fischer, J., Hong, J., Bentley, F.R., Munteanu, C., Hiniker, A., Tsai, J.Y., Ammari, T.: Panel: Voice Assistants, UX Design and Research. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. p. panel01. ACM (2018)
35. Kowalczuk, P.: Consumer acceptance of smart speakers: a mixed methods approach. Journal of Research in Interactive Marketing **12**(4), 418–431 (2018)
36. Krasnoff, B.: How to stop Google from keeping your voice recordings, vol. 2019 (2019), https://www.theverge.com/2019/5/13/18618156/how-to-stop-google-voice-recordings-storage-assistant
37. Lau, J., Zimmerman, B., Schaub, F.: Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. Proceedings of the ACM on Human-Computer Interaction **2**(CSCW), 102 (2018)
38. Law, E.L.C., Roto, V., Hassenzahl, M., Vermeeren, A.P., Kort, J.: Understanding, scoping and defining user experience: a survey approach. In: Proceedings of the

SIGCHI conference on human factors in computing systems. pp. 719–728. ACM (2009)
39. McCarthy, J., Wright, P.: Technology as experience. MIT press (2007)
40. McCormick, R.: Amazon gives up fight for Alexa's First Amendment rights after defendant hands over data, vol. 2019 (2017), https://www.theverge.com/2017/3/7/14839684/amazon-alexa-first-amendment-case
41. Moynihan, T.: Alexa and Google Home Record What You Say. But What Happens to That Data? (2016), https://www.wired.com/2016/12/alexa-and-google-record-your-voice/
42. Park, K., Kwak, C., Lee, J., Ahn, J.H.: The effect of platform characteristics on the adoption of smart speakers: Empirical evidence in South Korea. Telematics and Informatics **35**(8), 2118–2132 (2018)
43. Purington, A., Taft, J.G., Sannon, S., Bazarova, N.N., Taylor, S.H.: Alexa is my new BFF: social roles, user satisfaction, and personification of the amazon echo. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. pp. 2853–2859. ACM (2017)
44. Pyae, A., Joelsson, T.N.: Investigating the usability and user experiences of voice user interface: a case of Google home smart speaker. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct. pp. 127–131. ACM (2018)
45. Roto, V., Law, E., Vermeeren, A., Hoonhout, J.: User experience white paper: Bringing clarity to the concept of user experience. In: Dagstuhl Seminar on Demarcating User Experience. p. 12 (2011)
46. Seals, T.: Amazon Employees Given 'Broad Access' to Personal Alexa Info, vol. 2019. https://threatpost.com/amazon-employees-personal-alexa/144119/
47. Thompson, M.: Taking account of societal concerns about risk
48. Walsh, D., Downe, S.: Meta-synthesis method for qualitative research: a literature review. Journal of advanced nursing **50**(2), 204–211 (2005)
49. Wright, P., McCarthy, J., Meekison, L.: Making sense of experience. pp. 315–330. Funology 2, Springer (2018)
50. Wueest, C.: A guide to the security of voice-activated smart speakers (2017)
51. Yang, H., Yu, J., Zo, H., Choi, M.: User acceptance of wearable devices: An extended perspective of perceived value. Telematics and Informatics **33**(2), 256–269 (2016)
52. Zhang, G., Yan, C., Ji, X., Zhang, T., Zhang, T., Xu, W.: Dolphinattack: Inaudible voice commands. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. pp. 103–117. ACM (2017)