### Common TO DO

- : Make sure security/trust wording is not in section headers

### LOCKS THREAD

- RQ1: Kim
- RQ2: McKenzie
- RQ3: Mahsa
- Intro: Kim
- barriers Kim
- Reduce methods in study two to reflect updates to changes only to method Kim
- background related art: Kim

### STATUS

- Intro: Anita/Rakesh to do Boldification; September 22nd
- Background: Anita/Rakesh to do Boldification; September 22nd
- RQ1: Kim waiting on Anita/Rakesh for feedback and new direction
- RQ2: McKenzie waiting on Anita/Rakesh for feedback and new direction
- RQ3: Mahsa is waiting for Feedback
- Limitations/Threats to Validity: Kim to start a draft
- Related Work: Kim waiting on Anita/Rakesh for feedback and new direction
- Trust? : Kim will see if it would be interesting to bring it back

# Lights, Camera, Action: Do users understand the implications of home automation?

**Anonymous Author(s)**

## ABSTRACT

Home automation has become increasingly popular, with new interconnected products being introduced on a regular basis. While the benefits advertised by these devices have enticed end users, they may not fully understand the complexities of using these devices. Gaps in end users' understanding of how these devices operate, use and share their data (e.g., location tracking or automatically captured picture) can leave them vulnerable to security and privacy threats. In this paper we report on the results of a user study, where end users configured a set of three home automation devices (camera, light, lock). We found that participants had gaps in their mental models of how the devices operate, use, and share data. Further, the act of configuring these devices did not help participants correct their mental models. We also found that having a technical background did not always help.

## KEYWORDS

Mental Model Gaps, IoT, Home Automation, Security, Privacy

**ACM Reference format:**
Anonymous Author(s). 2017. Lights, Camera, Action: Do users understand the implications of home automation?. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems, Montreal, Canada, April 2018 (CHI'18),* 16 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 INTRODUCTION

[add preamble about growth of IoT] ***WHy care about IOT/Home automation: Lots of people are adopting it -> some marketing numbers
(NEED TO REVISE) We are now in the midst of a technology wave dubbed the *Internet-of-Things (IoT)*. Since Smartphones became mainstream about a decade ago with the arrival of the iPhone, the number of smart and connected devices (*e.g.,* tablet computers, e-readers, smartwatches, health monitoring devices, voice-controlled devices *etc.* ) has increased significantly.

This IoT technology wave is accelerating the proliferation of connected devices in homes. Smart door locks, thermostats, lights and switches that can be operated remotely using smartphones are just a few examples of popular smart home devices. Gartner predicts that future smart homes could have 500 or more smart and connected devices [**?** ]. In another recent report, Gartner forecast that the number of connected things will increase to 20.4 billion by 2020 from the 8.4 billion in 2017, with total spending reaching 2 trillion [**?** ]. While these numbers include all connected things, the consumer segment alone makes up 63% of the total. So what is driving this demand? A survey done in 2015 by iControl [**?** ], a software platform provider for major home automation companies (*e.g.,* ADT, XFINITY, AT&T, *etc.* ), found that *personal and family security* was the number one driver, closely followed by excitement about energy savings [**?** ].

***While people have studied configuration and impact of wi-fi and firewalls, etc. this is the first regular home owners, typically end-users are getting all these cool devices and are tasked with programming a network of devices on their own (problem). This is akin to how smart phones started whit early adopters and is now a staple. Refer to the gartner number.

***But this is a complicated situation «show figure»
***There are severe security/ privacy implications of EU are unaware of this complexity
So, now is the right time to understand how well end users can install and interact with these devices
*** OUr goal is to (multi-faceted approach to) understand how much EU do really understand about what these devices do
***We get towards this understanding via three directions:
***RQ1: we need to understand the mental models of the information flow because this will allow them to do the risk-benefit analysis of purchasing/ using the device
***RQ2: So even if they know how the data is flowing, do they realize who can access and control the device and the data
***RQ3: we need to understand whether they actually "get" what the device will do - by asking them about corner cases...because

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 ACM 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.

*CHI'18, April 2018, Montreal, Canada*
© 2017 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

if they dont understand what the device does and misconfigure the device then they will have security/ privacy implications.

*** we conducted a user study, where we recruited home owners who came to lab and configured a set of 3 devices (camera, lights, lock) and answered a set of questions.

***in broad strokes results

****3 types of mental model gaps- first time an eco system of networked devices. understanding about ..**

Our research is exploring the users mental models of home automation to better understand if these devices are comprehensible to home-owners. First understanding how users think their devices operate will identify gaps in their perceptions. Second to understand where they think the data, if any, resides and if that data is sensitive in nature. The third gap is determining if they can identify edge cases in their understanding of how the device operates.

Some previous work has explored users mental models of networking, including the Internet [5], understanding use of a learning thermostat [16], how users think about controlling access to home automation data [7], barriers to user ability to achieve security [1], perceptions of personal privacy in smart homes [14], and common mental model errors in feature interaction [17].

****Information flow why we should care**

Information flows or diagramming, as used in [5] were used to determine the users mental model of the Internet. Similarly we used diagramming to understand the users mental model of home automation devices. Home automation devices are not simple devices nor is their connectivity simple. By asking users to sketch a diagram on the information flow gives us insight to see if they understand the complexity and have comprehension of the interrelation with networking and the Internet.

****Point out why they should care using examples**

People purchase vacuum cleaners and may not know these devices collect, store, or share data. Not until recently was there a vacuum which does collect data and uses networking to share that data with the manufacture. What appeared as a smarter vacuum cleaner which doesn't run into walls also became a potential privacy issue when news posted the potential of the vendor selling or sharing users floor plans to retailers [15]. Had the users understood the data sharing it is unclear if they would still purchase or allow the data collection features. Had they drawn a flow diagram it is unknown how many would show the connectivity or were surprised by the news. Drawing the before, and again diagram after setting up the device, will show what their mental models was and if installing the device changed their mental model.

****Sensitive data sharing why we should care**

Many home automation devices send data across the public Internet including private information about the homeowner. Be it their personal video images or location by tracking their cellphone. As mentioned in [14] the perception of privacy is highly complex and involves different perspectives and dimensions. To understand their perceptions we asked participants a series of questions concerning data on the device and sensitivity of that data. Most of these home automation solutions require data and configuration information be stored outside the home. However the details are not shared with the user except for a brief pop-up when installing the app.

****The example of why we should care**

People have been buying and using Amazon Echo for years. It is unclear how many devices have been purchased. In fact 5 of our participants stated they have a voice controlled assistance at home and 2 mentioned the Amazon Echo. The Amazon Echo is always listening and when you say the key word Alexa it sends whatever you say to Amazon servers. Recently police asked Amazon to turn over all the voice recording for a Echo owner [11]. It is unclear how many purchasers knew Amazon stored and collected all this information. To understand their mental models we asked participants questions about the devices they set up and if the devices collected or shared private information.

****Configuration and operation why we should care**

People are installing home automation devices and setting up the different features example: scheduling the light to be on at a certain time. In [17]using a door lock example he asked participant to determine the status of the lock given certain scenarios. Understanding the edge cases of device operation is needed by users so they don't have their door open all day or their camera recording during improper times. Brush [1] in his study on home automation he found that users needed simple primitives so they could confidently configure and enable device. We checked users configuration success and their answer to scenarios as a way to understand their mental model.

****Example the Nest thermostat**

Nest thermostats are popular in the home to control home heating and air conditioning. Recent research, [16] found that users of the Nest device felt the device failed to understand their intent and the behavior of the device is hard to understand. He noted that systems struggle to understand human and human context and intent. If users find it difficult to understand how the device operates and are limited in their ability to program the device the actions of the device may be contrary to their intent.

## 2 RELATED WORK

**Mental Models:** have been adopted by researchers to study users' understanding of both security and privacy threats, and technologies (*e.g.,* [2, 8, 10, 13]). Friedman *et al.*'s study [2] of 72 participants on web security found that users from different backgrounds (rural, urban, and high-tech communities) had different security concerns. Wash's work [13] identified eight *folk models* of security threats that people rely on for making decisions about home computer security. Raja *et al.*[10] studied user mental models of Vista Firewall and explored ways to improve mental models without increasing the complexity of the interface. Klasnja *et al.*[6] undertook an exploratory study to examine non-technical users' understanding of Wi-Fi, their concerns with its use, and their practices to counter perceived threats. Similar to this work, Poole *et al.*[9] studied how end users deal with home network management, and what tools might help them with their task. However, their study was focused on technology of the day – home computers. More recently, Kang *et al.*[5] studied user mental models of the Internet and implications for privacy and security. Their findings suggested placing greater emphasis on systems and policies that protect privacy and security without relying on end-user security practices. Mazurek *et al.* [8] studied access control needs for home data sharing, and associated mental model gaps.

Other studies have examined either home automation or specific home automation devices. [17]studied users common mental models errors when examining feature interaction of a smart door lock. Ur *et al.*[12] studies how households use of home automation might impact teenagers. Yang *et al.*[16] studied the experience of people living with a smart thermostat. In contrast, our focus is in understanding the mental models (and gaps) that end-users may have when configuring and managing emerging smart home-automation devices; and the associated security and privacy concerns.

****OLD TEXT****

**Mental models:** are a commonly used methodology in psychology to elicit users' understanding about a problem or systems. Mental models are a model in a person's mind about how things work and help with decision making by enabling mental simulation of likely consequences of an action [2]. More formally, they are a "psychological representation of real, hypothetical or imaginary situations"[? ]. Mental models help us understand how users think about a system; understanding the gaps between end-user mental models and reality can help us identify pitfalls for system configuration or operation.

Mental models have been adopted by researchers to study users' understanding of both security and privacy threats, and technologies (*e.g.,* [2, 8, 10, 13]). Friedman *et al.*'s study [2] of 72 participants on web security found that users from different backgrounds (rural, urban, and high-tech communities) had different security concerns. Wash's work [13] identified eight *folk models* of security threats that people rely on for making decisions about home computer security. Raja *et al.*[10] studied user mental models of Vista Firewall and explored ways to improve mental models without increasing the complexity of the interface. Klasnja *et al.*[6] undertook an exploratory study to examine non-technical users' understanding of Wi-Fi, their concerns with its use, and their practices to counter perceived threats. Similar to this work, Poole *et al.*[9] studied how end users deal with home network management, and what tools might help them with their task. However, their study was focused on technology of the day – home computers. More recently, Kang *et al.*[? ] studied user mental models of the Internet and implications for privacy and security. Their findings suggested placing greater emphasis on systems and policies that protect privacy and security without relying on end-user security practices. Mazurek *et al.* [8] studied access control needs for home data sharing, and associated mental model gaps. In contrast, our focus is in understanding the mental models (and gaps) that end-users may have when installing, configuring, and managing emerging smart home-automation devices; and the associated security and privacy concerns.

**Barriers to Configuration:** In our study, barriers refer to the issues end-users have when configuring home automation devices. Barriers in computing are not new, prior research has focused on barriers in software development [? ? ] and to adoption of home automation technology [? ] for instance. Barriers that developers face have also been classified as either surmountable or insurmountable [? ]. As the name indicates, surmountable barriers are those that were eventually overcome, while insurmountable ones are those that couldn't be. We leverage this notion of barriers to articulate the configuration problems that participants in our study faced.

Ko *et al.* [? ] identified six barriers in developing software, namely, design, selection, coordination, use, understanding, and information. In our case, the use barrier is the most pertinent, since users may have trouble configuring and using a home automation device. Brush *et al.* [? ] identified four barriers to adoption of home automation technology, namely, cost, technology flexibility (or interoperability), poor manageability, and security concerns. Barriers related to poor manageability and security concerns of home automation are closely related to this work. However, since the work of Brush *et al.* [? ], home automation technology has undergone tremendous changes with the emergence of IoT, and the proliferation of smart phones and tablets that are used in configuring and controlling home automation devices. Our work focuses on barriers to configuring and setting up newer home automation devices that essentially claim to be plug-n-play.

**IoT and Home Automation Security:** Researchers have identified serious security and privacy concerns with smart home automation devices (*e.g.,* [? ? ? ? ? ? ? ? ]). A survey of the landscape of security and privacy concerns with smart-home devices and potential attack scenarios were laid out in [? ]. Similarly, Jose and Malekian [? ] surveyed different home-automation technologies from a security standpoint. Wilkowska *et al.* [? ] studied users' privacy concerns with smart home technologies, especially assisted living technologies. Oluwafemi *et al.*[? ] showed that even non-networked appliances (in this case CFL light bulbs) can be attacked through networked home automation systems. Ur *et al.*[? ] found that the smart home devices they studied did not have sufficient access control options. Defenses for threats against smart home and IoT devices and associated frameworks have also started to emerge. For instance, Fernandes *et al.*[? ? ] propose FlowFence that blocks undesired flows that are not specified by users in IoT applications. While those works focus on vulnerabilities in devices and protocols, our focus is on understanding end-users' perceptions of home automation devices, and implications for security and privacy.

## 3 STUDY1

In our first user study we performed an exploratory study to understand the barriers that homeowners face in actually setting up these devices. Participants faced multiple barriers, some of which were insurmountable. Our work indicates that current home automation devices run contrary to the perception that smart homes devices are “plug-and-play".

### Methodology

We recruited 7 participants: three female and four males; all within the age range of 35 to 64. A sample of 7 participants was chosen to ensure that each device would be configured at least 3 times. None had a computer science background.

The study was conducted in an IoT lab at Oregon State University. Participants were first asked to provide background information about their prior IoT experience, setting up networks, and on-line banking.

Then, the participant was asked a series of 5 questions. These questions were concerned with what the user thought the device did and how data related to the device was controlled and used. Additionally, they were asked to draw a diagram that indicated how data flowed between the app and the smart device in any form they choose. After the initial set of questions, the device was given to the participant in its box, reset to factory settings. They were allowed to choose between a provided Apple or Android tablet to install the device's app on. Participants then configured the device without assistance then stopped when they exceeded 20 minutes.

After each device installation or attempt, we asked the participant the same set of 5 questions that were asked before the configuration. However, in addition we also asked how their answers had changed since they had configured the device. We also asked if they wanted to make any changes to their drawing or create a new one.

### Results

We investigated three aspects in our exploratory study. First, we identified the barriers faced by participants when configuring the devices in section barriers. Barriers in correctly installing the device can lead to frustrations, and more importantly incorrect installations that might have security and privacy implications. Second, we find whether there are gaps in end users' expectations of the devices and how these devices manage their data in section mental models. A gap in an understanding how the device data can be accessed by others can lead to security/ privacy concerns.

### Barriers

Most participants faced barriers when they configured the devices. These barriers were in: (1) finding the right app from the App Store, (2) setting up the network, and (3) following the instructions provided in the device packaging.

Tables **??** and **??** show the number of steps that participants needed to configure each device. Green check marks indicate that participants were successful in the particular step, red triangle marks indicate a barrier that the participant was able to surmount, red stop sign indicates that the participant encountered an insurmountable barrier and stopped, and hollow black circles indicate steps that participants did not attempt either because they ran out of time or unwilling to go further. Dash in a cell is used to indicate that the participant did not configure that device. Note that despite Hue Light being a simple device, three (P1, P2, P5) out of the five participants failed to configure the device. The Nest Camera and Amazon Echo also had many barriers, which we discuss next.

***Finding the right app***. The most significant barrier was right in the first step – installing the app. Instructions provided in the device package asked participants to download the requisite app from from the App Store, but did not account for the presence of “imposter" apps or the app being missing in the App Store. For example, a search on “Hue Lights” provides the following apps (see Figure **??**). The first and many subsequent apps are third-party software for controlling Hue Lights. The second and fourth apps are Philips Hue, but one is from “Philips Lighting BV” and the other from “Philips Consumer Lifestyle”. Similar problems existed for other devices including Nest camera.

***Following instructions***.

Typical instructions provided in the pamphlet were deceptively simple, consisting of just two steps: plug-in the device and install the app. The instructions implied that the installation process is trivial. However, the opposite is true. In Table **??** and Table **??**, each of the symbols (check mark, triangle, stop-sign and hollow circle) represents one step in the installation process required to set up a device. There are 24 steps needed for the Nest Camera, and a minimum of 11 for the WeMo Switch. Such a vast discrepancy between the 2-step process implied by the instructions and the actual number of steps was a problem since participants started with a mindset that the process was going to be simple, and then were frustrated with the total number of steps that were actually required. P1 kept exclaiming the "*the directions are bad*" when setting up the Amazon Echo and WeMo Switch devices.

***Setting up the network.*** The Nest Camera and Amazon Echo had the most barriers (see Tables **??** and **??**). In the case of the Nest Camera, two out of the three participants faced barriers. For P1, the network failed three times when trying to configure, since she did not input the password for the network. However, this step was not made clear by the app, which provided no authentication prompts. Because of the Zero-Conf configuration style of Nest Camera, participants were not asked to input the password directly into the app. Instead it required them to move back and forth between screens in the app to do so. This caused confusion and added to the cognitive load of setting up the device. P3 had similar issues and gave up configuring the device after 4 attempts.

### Mental Model Gaps

To better understand mental model gaps, let us consider Hue Lights, which is a device that should have been relatively easy to understand and set up. However, every participant had gaps in their mental model about the device. An example of a gap is demonstrated by participant P1 (in Figure 1), where she thought that the iPad communicated directly with the light bulb. After installing the device P1 understood more about the dataflow, as she realized that the Hue Bridge communicated with the light bulb. However, she still had an incorrect mental model of where the data is stored (Figure 3).

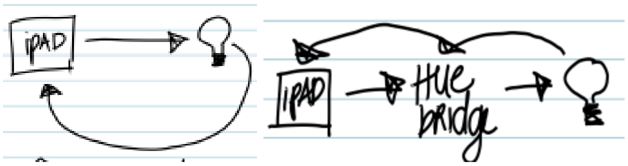


**Figure 1: Before setup** **Figure 2: After setup**

**Figure 3: P1's drawing of Hue Light communications**

P1 was not alone in having incorrect mental models even after installing the devices. On the contrary, many participants still had gaps after installing and using the devices. Most of the gaps continued to be in the use and control of data collected by the devices, as participants expected their data to be private and not accessible to third party vendors. For example, participant P2, who configured three devices (Amazon Echo, WeMo Switch, and the Hue Lights) never realized, even after connecting the devices to the LAN, that the data collected by these devices could be used by anyone else. On the contrary, she was convinced that *"it was the owner of the device who could only use the data"*

## 4 LESSONS LEARNED

### timebox vrs assist

Our original method included participants attempting to complete three devices in the 2 hours allotted. To accomplish the time-frame we limited each device install time to 20 minutes. We did not help the participant overcome any barriers hence, most devices were not fully configured or explored by the users. To determine if their mental models gaps are different or reduced the new study did not time box and offered participants assistance if they reached a barrier. We did this to see if the mental models gaps would significantly differ when they could complete all devices. To accomplish this we developed a list of potential barriers and solutions the researcher would help the participant if needed.

### NT to Technical

Our initial set of participants were local homeowners who answered a request for participants who were interested in home-automation. The homeowners who did participate did not have a technical backgrounds. Out process did not limit to either technical or non-technical we took took them as they came. Not including a mix of backgrounds our exploratory research did not potentially include the diversity of people who would be installing home automation devices. To examine the potential differences of technical background for barriers and gaps we decided to expand our research to incorporate a 50/50 mix.

### 6 devices to 3

Number of devices - six was too many kept camera lights and locks, Alexa was just a layer on top. From which we selected a subset. Chnages of devices-

All people all devices

Scenarios - if they understand what the configurations mean therefor sceanios.

## 5 STUDY 2 METHODS

Need to add the correct flow and identify what are the key parts that we need to an eye out

****We conducted a user study because we were interested in understanding if the end users understood securing and privacy implications of using home automation devices.** We conducted a second user study to investigate how our target population configured home automation devices. Specifically, we were interested in evaluating the usability of these devices, their security and privacy configurations, and whether end users understood the security and privacy implications of their configurations. Our target population was working individuals 18 or older who were homeowners, because this population is likely the most typical consumer of smart home devices.

## Participants

We recruited a total of 24 participants (See Table 1) of which 12 participants with technical background and only 8 having a degree in computer science, along with 12 non-technical. Technical background was based on prior educational areas in the engineering department. To understand their background we asked about prior IoT experience. Table1 shows the breakdown of the participants in age, technical background,computer science background and how well they answered technical questions. To recruit participants, we requested an advisor for the Department of Electrical Engineering and Computer Science (ECCS) at <Anonymized> University to send out an email to the ECCS faculty and student mailing list. We also posted flyers around the college calling for interested participants. In addition, we posted recruitment ads on local Craigslist, Nextdoor.com and Facebook pages. We required participants to be homeowners and have previous experience with installing apps on an iPhone or Android device.

To understand participants educational backgrounds, we asked participants to report the major of their highest degree. Also, to understand participants' previous IoT experience regardless of their educational background, we asked them technical questions to understand their background knowledge on privacy and security.

## Study Design

****People get introduced to devices and were asked to draw flow diagrams. We chose these devices because of their popularity**

**Table 1: Study Participants (Total=24)**

| Participants | Gender | Age | Education | Technical Questions that were Correct (out of 5) |
|---|---|---|---|---|
| *Technical Participants* | | | | |
| P1 | M | 33 | CS PhD | 3 |
| P2 | M | 58 | CS | 4 |
| P3 | F | 32 | CS | 3 |
| P4 | M | 23 | CS | 2 |
| P5 | F | 56 | CS | 2 |
| P7 | F | 23 | CS | 0 |
| P8 | M | 69 | CS | 3 |
| P9 | M | 24 | Math and Science | 4 |
| P10 | M | 74 | Engineering | 0 |
| P13 | M | 27 | Electrical Engineering | 4 |
| P16 | M | 39 | CS | 2 |
| P23 | F | 23 | Chemical Engineering | 0 |
| *Non-Technical Participants* | | | | |
| P6 | F | 19 | New Media Communication | 2 |
| P11 | F | 61 | Accounting | 2 |
| P12 | F | 68 | French | 2 |
| P14 | M | 50 | Communications | 0 |
| P15 | M | 40 | Business Administration | 3 |
| P17 | F | 20 | Bioenergy | 1 |
| P18 | M | 49 | Material Sciences | 1 |
| P19 | F | 61 | Physical Education | 3 |
| P20 | M | 70 | Aviation Maintenance | 2 |
| P21 | F | 52 | PhD Education | 4 |
| P22 | F | 39 | Geology | 2 |
| P24 | F | 44 | Public Health | 0 |

*Device Introduction and Scenario Setting:* Participants were introduced to the three devices and their functions as shown in Table 2. These devices were picked because of their popularity. Specifically, smart cameras, smart lighting, and smart locks were reported to be some of the smart home devices that people are ready to spend on We first asked the participants to "please draw a diagram visualizing the expected information flow, to- and from- the device." To lessen the learning curve, we gave them two warm-up tasks:

- Change the network setting to the local WiFi router
- Add a new free game using the app store

We asked participants to talk aloud during the warm-up tasks and provide feedback on how they were completing the task.

****Then we gave them scenarios for setting up each device.** After they completed the flow diagrams we then gave them the scenarios for setting up each device (See Table 3). Our purpose with the scenarios was to give each participant

Table 2: Devices and the description provided to participants

| Devices | Description Given to Participants |
|---|---|
| Nest Indoor Camera | The Nest Cam Indoor security camera is designed to help you look after your home and family, even when you are away. 24/7 live streaming so you can check in anytime and Two-way audio to hear what is happening through microphone. |
| Hue Light Kit | Philips Hue combines brilliant and energy-efficient LED light with intuitive technology. Together, the light, the bridge and the smart controls will forever change the way you control and experience light. Works with Alexa. |
| Insteon Lock Controller | With a Lock Controller, you can control access to your home from any Insteon device in your home. And if you're not at home, you can use your smartphone, tablet or Apple Watch to lock and unlock your doors when paired with the Insteon Hub. |

Table 3: Scenarios given to assist device configuration

| Scenarios to Configure Devices |
|---|
| You typically leave your house during the weekdays in the morning and arrive back in the evenings.,To save energy you turn the lights off and set the camera to be on when you are gone. |
| During the weekdays, M-F, you set you lock to unlock when you typically come home. You schedule the lock to be locked after you leave. |
| On the weekends, you usually get up later and stay home and set the camera to be on only at night when you go to bed. Your lights are not scheduled on weekends. |
| Additionally, when you are home you set up the camera and lights to detect your presence by turning off camera and turning on the light. |

a typical task a homeowner might come across when setting up the devices for themselves. Tasks included scheduling the lights, camera, and door lock for specific time settings.

****Participants were asked questions to get initial understanding of how they saw data on the device** Participants were presented with the first of three devices and then asked the same questions in exploratory study (See Table 4).

*Set up device.* ****After the questions, we asked them to set up the devices without assistance but researcher stepped in occasionally.** After the initial questions, the participants were asked to setup the device and configure the scenarios. They were given an Apple iPad or an Android tablet to install the app on. The participant then configured the device without assistance unless they hit a barrier and their progress was stalled. Researchers encouraged the participants to keep trying and after 3 failed attempts by the participant or 2 minutes without resolving the barrier.

Table 4: Questions asked before and after device configuration

| Questions | When were the Questions Asked? |
|---|---|
| In what ways, do you expect this device to meet your day to day needs? | B/A [i] |
| Who can operate the device? | B/A |
| Is there data on the device and what purpose is it used for? | B/A |
| Does this device contain any information and you would consider sensitive? | A [ii] |
| What kind, be specific? | A |
| Does this device share sensitive information with anyone but you? | A |

[i] B/A = Before & After Scenarios [ii] A = Only After Scenarios

### **Final questions

*Post configuration:* ****After they finished the scenarios we asked them same privacy questions and other data sharing questions. Also, asked them to redraw mental model.** After the participants were done configuring the device, to see how the participants mental models changed, we asked the participants the same questions regarding privacy. Also, we asked them to redraw or change their flow diagram based on their understanding of the device after configuring the devices. In addition, we asked them several other questions as shown in Table 4. We also asked how their information might be compromised and if they had any privacy concerns and to rate the concern on a scale of 1-5.

*End of study questions:* ****At the end of the study we checked to see if devices were configured correctly and what may have gone wrong. We also asked new scenarios** After the participant attempted to configure all three devices and the post questions were answered, we checked the participant's configuration of the devices based on the scenarios we gave them. If a device was configured incorrectly or not completed, we asked the participant to explain why they did not complete the configuration as asked to ascertain if it was an user error or complications with the device configuration.

****Then we gave them 5 additional scenarios to ascertain if they understood the implications of the configurations** In addition, we asked them 5 new situations/scenarios (as shown in Table 5) to ascertain if they understood

**Table 5: Scenarios given to participants after device configuration to ascertain their understanding of the device**

| Post Configuration Scenarios | Correct Answers |
|---|---|
| During a weekday.,You drive home and park on the street. Your car is full of groceries and other shopping, which take many trips to bring into the house. Five minutes after you drove in, you are still making trips to the car. Is the Camera on or off? | Camera: On |
| During weekdays,You come home early instead of the normal hours, are your lights on when you come in? Is your door unlocked? | Lights: On, Lock: Unlocked |
| You host a July 4th, barbecue party on Monday in the front yard, is your camera recording? Also, a neighbor tried to enter your house, will the door unlock? | Camera: Off, Lock: Unlocked |
| You left your phone in your car when you come home during the week, is your door unlocked and is your camera on when you come inside? | Lock: Locked, Camera: On |
| You left your phone at home when you left for the day. Your friend enters the house will the camera be off? | Camera: Off |

the implications of the configuration. For example, for the first scenario in Table 5 the correct answer is a Camera should be "On" because the camera is still on and recording daily activities.

### Data Collection:

**We collected all this data by audio recording and transcribing.** All the sessions were transcribed by the researcher. In addition, the participants were audio recorded and screen captured with Quicktime video recording of the iPad while they used the devices and set up the apps.

### Data Analysis:

We performed qualitative data-driven analysis by using the method of card sorting on participants' flow diagrams to answer our first research question. We began with open coding to identify categories of mental model gaps and then continued sorting diagrams into these respective categories. This analysis was performed using negotiated agreement. We also employed the ANOVA statistical test to determine significance between certain factors in our data.

Need to mention how categories were identified in RQ1. And IRR for other things in RQ2, RQ3

[I cover the categories in RQ1 but check make sure OK. It's more results than the methods.]

## 6 STUDY 2 RESULTS

Our analysis shows homeowners are not cognizant of the complexities and risks home automation can give them. Our participant data shows people do not understand these devices are as connected' as their phones. Participants continually placed these devices as unsophisticated and less private than their smartphones. Home owners are aware of privacy on their smartphones but this awareness is not transferred to appliances in the home. Lack of awareness can lead to data loss, hacking, and physical security lapse. In the following sections we discuss what participants know and believes these devices do and how that is contrary to their benefit.

### RQ1:What are the Information Flow Mental Model Gaps and Implications?

The key goal of RQ1 was to analyze whether homeowners are able to understand the connectivity when adding these devices into the home. First, whether they could accurately draw or describe how these devices communicate inside and outside the home. A gap is when their understanding how these devices communicate is either wrong or missing key components, especially outside communication. Second, if after installing these devices they still had a gap or were not able to understand the correct data flow.

** **People make mistakes**

*Flow Diagram Mistakes.* We had the participants draw a diagram visualizing the expected information flow, to and from the device. We used their information flow diagrams from each device to understand participants' mental model and their perception of how the device communicates. We asked participants to draw an information flow diagram for all the devices before they attempted installing them. Then after each device installation we asked them to accept their past drawing or make a new drawing. We had 8 out of 24 participants who did not provide describable flows mainly because they did not understand how to draw an information flow diagram or flow chart. They provided more of a text describing how they would use the device not how the device communicates. Analyzing the rest, we found that many participants' information flows were not correct.

****To understand why they happened we categorized them**

In order to better understand the mistakes we categorized them using card sorting revision by multiple researchers. Reviewing their flows it was easy to see which areas they

**Table 6: Description of categories used to identify mistakes**

| Categories | Description |
|---|---|
| Missing Data Going Outside, Missing Data Stored Outside, Indescribable Flow, | Missing hubs or bridges, router, Internet, or cloud to show data is going outside the home. |
| | Missing that the data is stored on an external cloud service or server account. |
| | Something other than a flow diagram or the flow diagram was not readable. |

**Table 7: Flow Mistakes by Device**

| // | Camera | Light | Lock |
|---|---|---|---|
| Data Going Outside [iii] | 14 | 35 | 29 |
| Data Stored Outside [iv] | 9 | 13 | 12 |
| Indescribable Flow [v] | 10 | 9 | 11 |

[iii] Data Going Outside = Missing hubs or bridges, router, internet, or cloud to show data is going outside the home [iv] Data Stored Outside = Missing that the data is stored on an external cloud service or server account. [v] Indescribable = Same as before

**Table 8: Number of Corrected Flows by Device**

| Added After Changes | Camera | Light | Lock |
|---|---|---|---|
| Data Going Outside [vi] | 0 | 3 | 5 |
| Data Stored Outside [vii] | 3 | 0 | 0 |
| Indescribable Flow [viii] | 5 | 1 | 0 |

[vi] Data Going Outside = Missing hubs or bridges, router, internet, or cloud to show data is going outside the home [vii] Data Stored Outside = Missing that the data is stored on an external cloud service or server account. [viii] Indescribable = Same as before

missed when compared to the expected information flow (See Figure 4) similar to Authors [3] where they provided an architecture diagram for smart homes which closely matched the flows of our devices.

The identified categories are: 1. Missing that data is going outside the home. 2. Missing that data is being stored outside the home. 3.Indescribable flow. (See Table 6) which shows the the categories and a description of each category. See Table 7 for all mistakes by device.

** **Mistakes persisted after device installation**

Mistakes persisted after device configuration. While it is understandable that participant flow diagrams prior to device installation are incorrect, flow diagrams after installing and configuring the devices did not correct their gap to any greater amount. (See Table 8).

** **Two categories**

The most frequent gaps participants made is not understanding data goes outside the home. We collected the data by the number of mistakes by category and by device. (See Table 7) with the categories and the number of mistakes per device.

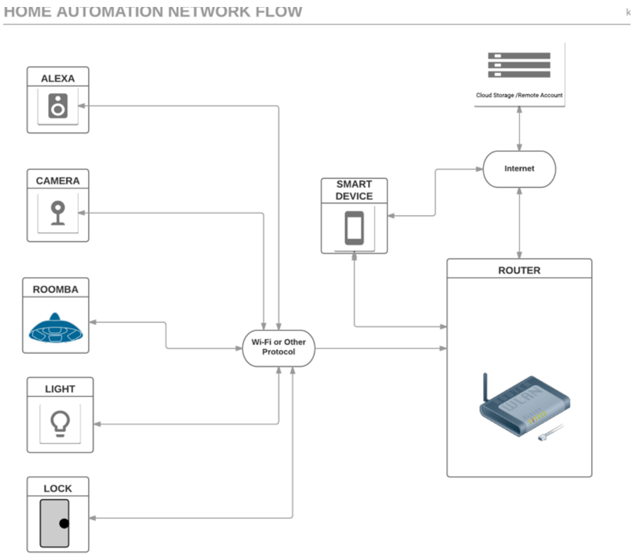


**Figure 4: Home Automation Network Architecture**

** **Missing data going outside the home had the most mistakes even after the device warned them data was being collected**

The category with the most mistakes was data going outside the home. Participants who missed this area usually drew a diagram with a direct connection from the iPad to the device. Some included an internal network but their gap was not showing the data went outside the home using a router or Internet connection. Eight participants, (See Table 8), after configuring the device did add to their diagram an external connection. Most improvements in gaps was with the Lock by participants adding hubs or Internet to their flow. Prior to installing the lock it did not appear to have any other connection except the iPad. After installing they used the Insteon Hub which was directly connected to the router and the lock controller with the lock.

The next categories of mistakes was users missing data being stored outside the home. A server login is required for each device. Participants as part of the setup needed to log into the account for each device. Once logged in they received a message to accept the vendor's use of the data or locations data. Data stored at the server might be schedules as in the camera and light or usage information or video images in the case of the camera. The camera and light also collect location data and both camera and light displayed an affirmative warning that required a positive selection by the user to accept. However, majority of participants did not include in the flow diagrams, from before or after

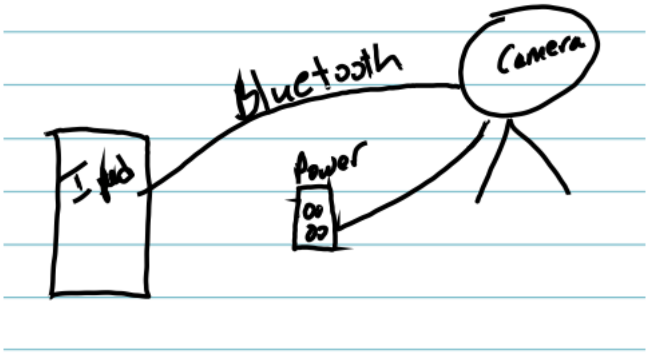


Figure 5: P11 flow diagram of the Nest camera

device installation, the presence of an external server or an account on the server. Data shows, (See Table 8), where there were only 2 participants who after configuring showed or described the external cloud or server.

** **Example of missing both**

As an example, in P11's (non-technical) drawing, (See Figure 5), she knew there was some communication method to the camera, albeit shown incorrectly as Bluetooth. However, she did not include the Internet or the server account at Nest and not even a local router. P11's after image was essentially missing the same components and when asked why they drew it this way she said: "Because the iPad is with me and I can see what's going on at home.' She meant the video images she was looking at on the iPad when she said this. The gap in her mental model was that she failed to understand that the iPad actually communicated to the camera over a wireless LAN where the images are transmitted via the Internet and stored in the cloud. See (See Figure 4), which shows the smart device connecting to the cloud servers, as was in this case.

****Next we look at the stats we see no difference between technical and non-technical**

*Technical vs. Non-Technical.* Our participants included both technical persons with engineering degrees or persons with non-technical degrees or experience. To see if there were any differences in articulation of device flows among participants with and without technical backgrounds, we counted all the flow mistakes then ran an Anova repeated measure between the technical and non-technical participants. The results are shown in (Figure 6)and (Figure **??**).

The results suggest there is no significant difference in average number of gaps (See Figure 6), when counted across devices, across technical, and non-technical participants.

****Light we found difference between T&NT, where NT were worse**

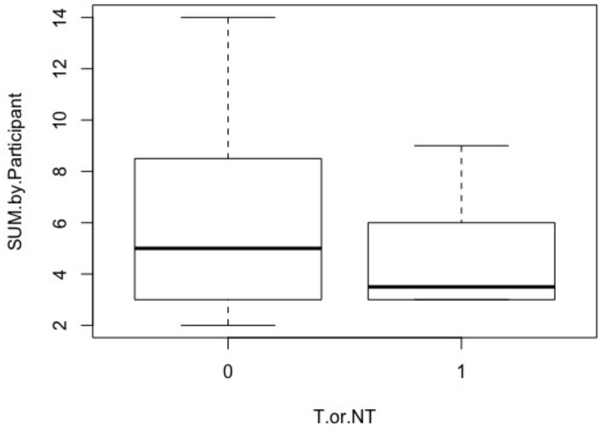


Figure 6: Diagramming comparison between technical and non-technical

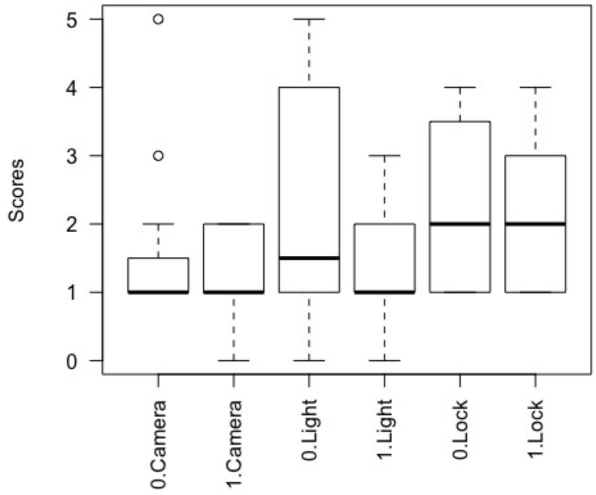


Figure 7: Diagramming comparison between technical and non-technical by device

However, looking at the results per device we found differences in the number of gaps across technical and non-technical participants. (See Figures 7) and (Figure **??**).In particular, looking at the differences between technical and non-technical participants for the Hue Light, non-technical participants did better on average although there is more variability in their performance and this difference was found to be significant.

** **Camera was easiest ref to means followed by light and lock**

Results also show that with the camera, people performed similarly between technical and non-technical participants (See Figures 7). This can be explained, in part, by the fact that the camera did not have any additional hardware as the other devices and was easier to install, (*e.g.,* light had lock). Further, six of our participants mentioned they have video cameras at home and it was split between the two types of background.

** **Connect mistakes to potential security and privacy threats**

The flows made by participants, even the technical skilled users, made lots of mistakes. A lack of understanding of the networking and external connection can lead to multiple risks the user may not even be aware of. The list below explores some of those potential risks:

** **I know where you live**

1. Localization and tracking of the person's movements and whereabouts. This was seen in the Nest camera and Hue lights. They allowed one to easily set up the camera and lights to conveniently turn off or on when home or about. However, they do not give any sufficient warning to the installer that this data is kept remotely and is collected whenever you leave or come home. It is unclear if they continuously keep track of your movements and to what detail. A simple pop-up did note that location information was being used but we found all participants moved quickly and selected yes. And as was demonstrated by their flow diagrams, participants made mistakes when it came to external connections, Internet, cloud, or server.

** **I know who you are**

2. Identification of person either by name or face. The camera collect still image photos whenever the camera notices movement. The image is collected and sent via email to the user or app along with the images and video kept at the cloud server. The images are stored until you manually go in and delete them. Information collected by the devices can uniquely identify a person by face or address. Address information is collected for the camera and light so they can track when you enter or leave your home. Many of the devices allow you to add family to the device to see or control the camera or light leading to a connection of familial data.

** **I know what you did**

3. Profiling an individual or household. Profiling requires data from multiple sources which allows hackers to combine multiple sources to get a more complete picture. IoT devices are usually single devices coexisting amongst other devices. However, more devices and software is available which allows you to manage and operate dissimilar devices and technology as one. Amazon Echo was owned by a couple of participants, they mentioned and showed in their data flow the fact they could control the light with Alexa. However, usage and information about the devices is aggregated in multiple hosting sites either Amazon or other aggregation data warehouses. Profiling them becomes easier to the unsuspecting homeowner.

**Table 9: Answers to data questions**

| | Camera | Light | Lock |
|---|---|---|---|
| Data On Device | 20-yes | 16-yes | 13-yes |
| Sensitive Data | 18-yes | 13-yes | 17-yes |
| Shared Data | 10-yes | 8-yes | 7-yes |

**Table 10: Categories of data**

| | Camera | Light | Lock |
|---|---|---|---|
| Phone Location | 2 | 6 | 7 |
| Image Video | 9 | 0 | 1 |
| Schedule | 1 | 1 | 2 |
| Personal | 4 | 4 | 15 |
| Usage | 0 | 0 | 2 |

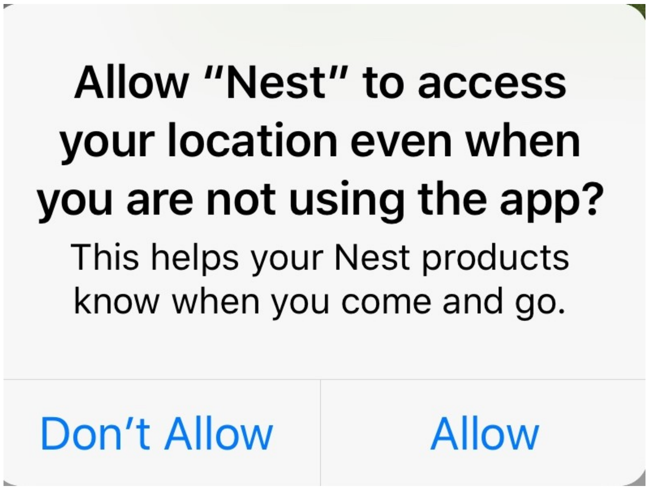
## RQ2:What are users mental model gaps with data?

A lack of understanding of how and where data is stored can lead to data leaks and misuse of personal data. We wanted to see whether gaps existed and to what extent. Therefore, we asked a set of 6 questions (see method, table 4 ) asking about data usage, data sensitivity and sharing data. We asked them these questions both before they set up the device and after. But their answers were similar, so here we report only the after configuration data. Along with not seeing a difference in answers of technical or nontechnical participants.

Table 9 presents the answers to the participants' perceptions about whether: (1) there is data on the device, (2) if that data is sensitive, and (3) if that data is shared.

We found participants had gaps when they did not realize their data on devices was sensitive. Especially when they knew the data on their phones was sensitive (23 out of 24 said yes to sensitive data being on phone). Through our findings we found that users do not correlate the sensitive data on their phone to their IoT devices. However, most sensitive data on the phone is local to the phone but not with the home automation devices where data is remote on cloud servers. Thus leaving the gap of users understanding of data. We now look at answers from our study for each type of device(Camera,Light and Lock).

*Camera*:Most people got the first two questions correct regarding data on the device and if the data was sensitive (See table 9). This might be because people are already sensitized to video/images being captured and used in their phones. For example, P14(Not Technical) realized that the sensitive data can be misused: "*[Video content that is stored on the camera] can be used as blackmail*"



Allow “Nest” to access your location even when you are not using the app?

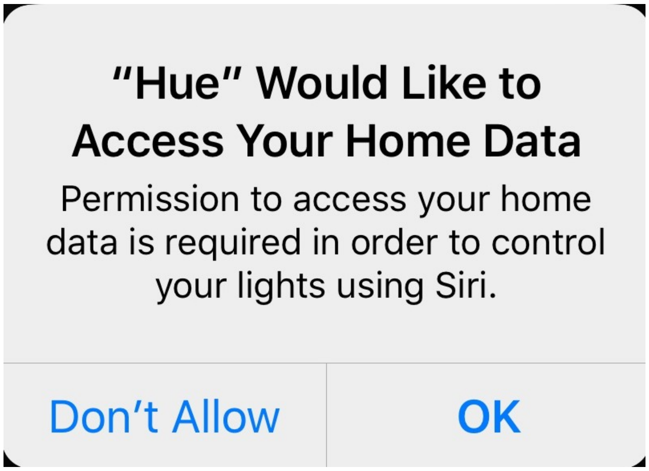
This helps your Nest products know when you come and go.

**Figure 8: Notification from camera app**

What was counter-intuitive was that despite them saying that there is sensitive data on their devices, only less than half (See Table 9) of (10) participants said yes to sharing sensitive data. This is even more concerning, since during the configuration the camera explicitly asked user permission/ provided notification that Nest would collect information (See Figure 8). Participants either did not pay attention to the pop-up or did not make the connection that their data could be shared with the vendor. P6(Not Technical) said“*Doesn’t appear so,not sure* ”. The gap in this area is users are agreeing to terms, but not making the connection of what the agreement states and the impact it has on their data sharing.

*Light*: The largest gap was with the light, less people did not realize that there was data stored on the device(33%) or that it is sensitive(45%). The light stores its location and the device location that is controlling it(smart phone, tablet, computer,etc). Information is also stored about users light schedule, which could allude when the user is home or away.People had the impression that this is just a light and it can be unplugged easily if hacked, as noted by participant P9(Technical) “*it’s just lights*”. As a consequence, even fewer people (33%) realized the sensitive data can be shared with others(Company, other users on account,etc). This leaves the users at risk if this information falls into the wrong hands.

*Lock*: A majority of the participants (See Table 9) assumed that the lock is storing data, which they marked as sensitive. In the case of the lock, the data is actually not stored on the device or shared. Therefore, about 70% people got it wrong. This might be a case where participants just assumed the lock stores data, as revealed by P12(Not technical): “*[Data on the device purpose] knows how to turn on and off lock. Knows date and time. Knows what the on and off positions are.*”. Following the same trend 7 participants incorrectly assumed that their sensitive data was shared.



“Hue” Would Like to Access Your Home Data

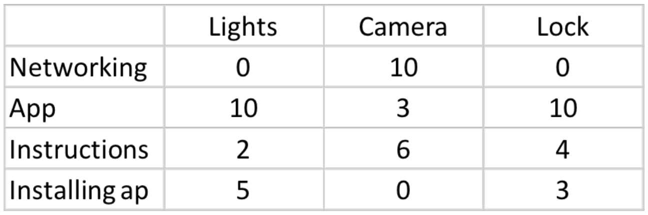
Permission to access your home data is required in order to control your lights using Siri.

**Figure 9: Notification from light app**

When looking across these devices, we see a consistent pattern where few participants realize that the data can be stored with the vendor. Only 2 participants (P21, P5) mentioned needing to look deeper into the privacy agreement; P21(Not Technical) “*Didn’t read or see that information would have to read and see*”.P5(Technical) notes that “*must read user agreement*”.This highlights that users don’t read privacy agreements nor think about them. A few recent examples, one is the case with the Roomba vacuum,[15], the device has a sensor to detect walls, steps,etc. What the device was actually doing was making floor plans of users homes and those floor plans will now be shared with companies like Google and Amazon to advertise to those users. The second example is the gap users not reading privacy agreements. The Amazon Alexa stores voice history and vocal commands that have been said to it (ex: Alexa what is the weather, Alexa play country music) there is a record for all the statement that is said to the device. Which has been used previously in a court of law.

Our devices could store the following categories of data (See Table 10). We wanted to investigate whether our participants realize what kind of data is actually being stored, as this affects the kind of vulnerability that participants face. For example, in the case of the camera (as well as other IOT devices) it stores the locations of the camera itself and the location of the device (computer, smart phone, tablet, etc) which is used to control the camera. This leaves the users at risk of break-ins if this information falls into the wrong hands. As well as having two way radio that can be controlled from the app.

Table 10 shows that very few participants were aware of the different types of data collected. Looking specifically



| | Lights | Camera | Lock |
|---|---|---|---|
| Networking | 0 | 10 | 0 |
| App | 10 | 3 | 10 |
| Instructions | 2 | 6 | 4 |
| Installing ap | 5 | 0 | 3 |

**Figure 10: Number of barriers per device**

at phone location very few people were right (See table 10 camera:2, Light: 6 ). This was after they had configured the devices to allow location access, where light and camera asked users more than once to access location information. See figures ( 8, 9 ). What was concerning was that only 37% of participants said that the camera stored video images, when the device itself is a video camera that stores content for thirty days. As well as the ability to watch live action video.

In summary, participants have gaps in their understanding of the data being stored and shared. Very few people recognize the kinds of sensitive data (schedules, location, etc) that are being continuously accessed and logged, despite them affirmatively given such access to the device. This implies that people might just be giving access without recognizing what it means to actually give permission to their data. Meaning that they are not thinking about the implications of someone having access to the data. In A recent study done by Harkous *et al.*[4], they found that these devices " *request more permissions than they actually need*". This is consistent with our results where people didn't understand where their data resides and extent of their data being collected.

### RQ3:What are the barriers users face when configuring these devices?

Most participants faced barriers in setting up a device. Participants were allowed to fail and try to find the solution on their own. If they were not able to continue the configuration the researcher helped them to get over the barrier. Participants were allowed to find the solution but, after 2-3 minutes they were assisted by the researcher. Otherwise, the researcher was just observing and taking notes about where they had difficulty. If the participant needed assistance, researcher helped the participant move forward to the next step in configuration.

Many of the barriers faced by our exploratory research, connecting the device to the network, installing the right app, following the instructions, and configuring the device through the app. (See Figure 10) ,shows the participants and the total number of barriers they faced throughout setting up three devices. Every participant faced a barrier setting up the devices (See Figure 11). When comparing the technical and non-technical there was not significant difference in the amount of barriers faced by either group.

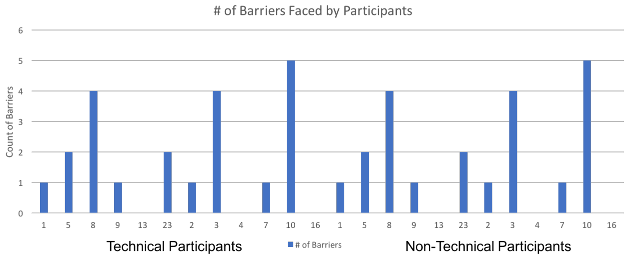


**Figure 11: List of participants and the number of barriers they faced**

The biggest barriers participant faced was using the app to configure the devices. When using the app to configure either the Lights or Camera participants faced the most barriers. With the Hue Lights participants were given a schedule to setup the lights to come on and off, see methods section. Participants either could not find where in the app to add a schedule or they had issues with the complexity of the app to set and save the schedule. With the lock participants hit barriers when trying to add the lock within the app.

### RQ3:What Mental models gaps users face after configuration was able/not able to answer scenarios?

Motivate the use of scenarios – how well participants understand the implications of their configurations.

It is important that users of home automation have a strong understanding of how devices are configured and work. It would benefit them to be cautious about the cases that could put them at risk. For example, users should be careful when scheduling their camera to be turned on or off to ensure that when they are away from home, they can monitor the house. To understand how much users think they know after configuring the devices, we provided participants with five different scenarios involving eight questions about the devices. These scenarios include everyday activities and special events that users would be involved in. Table **??** shows the scenarios and the devices that are mentioned in each. We designed these scenarios to explore how well participants understood the implications of configurations they did.

introducing one example scenario in detail - phone left at home.

In each scenario, we asked participants to determine whether the state of the mentioned device would be on or off. Participants answered these questions based on how they configured each device. In one scenario, participants were asked to state the status of the camera when they left their phone at

**Table 11: Devices involved with Each Scenario**

| Scenario Name | Lock | Camera | Lights |
|---|---|---|---|
| Grocery shopping trip | - | SQ1 | - |
| Come home early | SQ2 | - | SQ2 |
| 4th of July barbecue | SQ3 | SQ3 | - |
| Phone left in car | SQ4 | SQ4 | - |
| Phone left at home | - | SQ5 | - |

**Table 12: Incorrect Answers to Scenario Questions**

| | Lock | Camera | Lights |
|---|---|---|---|
| Technical | 15 (41.67%) | 25 (52.08%) | 3 (25%) |
| Non-technical | 12 (33.33%) | 23 (47.92%) | 5 (41.67%) |

home for the day and somebody entered their house. We call this scenario "phone left at home". According to the device configuration, the camera should be turned off. In this scenario, four technical participants got the correct answer and also correctly configured the camera while this number for non-technical people was only two. Moreover, the number of non technical participants who correctly configured the camera but answered the scenario incorrectly was four. This shows that people did not necessarily have a solid understanding of device functionality even when they configured it correctly.

[ kk-I would give the example in quotes on a separate line like Yarosh did. ]

Results 1: Let us look at the results of answers to scenarios. Table of averages and T/NT

The result of participants' answers to scenario questions is shown in Table ? per device. We found that participants with a technical background answered 4.33 scenarios correctly out of 8 while non-technical participants answered 4.67 correctly.

$^{**}$ **results overview**

As shown in Table **??**, we found that many participants answered scenario questions incorrectly. As Table **??** shows, the percentage of incorrect answers by technical people is even more than percentage of incorrect answers by non-technical people for the camera and lock. This result is interesting to us to explore more about the main effect of participants' background on their answers to scenario questions.

$^{**}$ **explore dependency of number of correct answers on users' background as T and NT**

To explore a little more in depth on the relation between technical background and correct answers, we used the Anova test. As we have three devices, to control the effect of devices on right answers we used blocking methods. The result shows that the interaction effect is not statistically significant between device and background. $F(3, 106) = 2.086, p > 0.05$. In other words, the relation between correct answer and being technical or non-technical does not depend on type of device. Anova also suggested that the main effect of background and type of device is not statistically significant.($F(2, 107) = 0.016, p > 0.05$ ;$F(2, 107) = 2.364, p > 0.05$ in order)

$^*$**explore dependency of right answers and on user background per device**$^*$

$^*$**explore dependency of configuration on answers**

As we got that there is no interaction between device and participants' answers, we did chi square test to explore dependency between configuring correctly or incorrectly and giving answer right or wrong. Chi square test suggested that relation between right and wrong answers and configuring correctly or incorrectly is not statistically significant. the result for camera and lock is as follows: ($\chi^2(87, 2) = 1.3803, p > 0.05$); ($\chi^2(86, 2) = 1.9115, p > 0.05$) In other words, doing configuration correctly would not lead to have a correct mental model about implication of configuration.

security and privacy implications of this

The results show people even with technical background would not have a well mental model about the configuration they have done before. It shows the threat which having technical background would not be enough to be safe of privacy threats. Moreover, configuring correctly would not guarantee that users have correct perception about configuration implications. These results would motivate vendors to provide some serious approaches to increase security and privacy awareness of people about the implication of smart home automation functions.

## 7 IMPLICATIONS

In RQ1 the second largest number of mistakes was with users not showing that data moved externally outside the home. As part of the install users used a provided email account and password. The threat in our results are the participants did not use their own personal account and by providing them an account they may not of perceived the account used the outside server to login.

Another threat to both RQ1 and RQ2 is participants came to install the devices in a closed environment, not their home. Seeing pop-ups and quickly dismissing the pop-up during the install may have contributed to their lack of understanding that the pop-up mentioned location data or stored data.

Both of these threats may change some of the results however, many of the participants were correct missed other areas consistent with our results.

## 8 CONCLUSIONS

Homeowners wising to integrate home automation into their environment need to understand the complicated interaction between the data stored on cloud servers and their data privacy. Homeowners need to learn the initial steps for configuration and also learn how and what data is being collected then stored. Preferably their should be an opt-out feature in all devices, as is incorporated in the Hue lights, which allows you to locally control lights without the Internet. We saw that users who were given minimal pop-ups about location tracking or data storage did not pay attention and always hit yes.

Our study showed participants had gaps in all areas of our research, from knowing where their data resided, understanding that data is collected, the type of data collected, to understanding the state of the device when acting outside the configuration. All of which, being unaware could lead to privacy and security concerns. Because of this, privacy-by-default should be incorporated in all settings which protect the homeowner.

Technical background did not directly improve homeowners ability to explain how devices communicated, understand where data resides, or being more sensitive to data privacy. Our work suggest a need for further research into privacy protection mechanisms that forewarn users and gives them the ability to modify data collection without losing functionality.

## ACKNOWLEDGMENTS

TBD

## REFERENCES

[1] A.J. Bernheim Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, Stefan Saroiu, and Colin Dixon. 2011. Home Automation in the Wild: Challenges and Opportunities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 2115–2124. http://doi.acm.org/10.1145/1978942.1979249

[2] Batya Friedman, David Hurley, Daniel C. Howe, Helen Nissenbaum, and Edward Felten. 2002. Users' Conceptions of Risks and Harms on the Web: A Comparative Study. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems (CHI EA '02)*. ACM, New York, NY, USA, 614–615. https://doi.org/10.1145/506443.506510

[3] Dimitris Geneiatakis. 2017. Security and privacy issues for an IoT based smart home. https://doi.org/10.23919/MIPRO.2017.7973622

[4] Hamza Harkous, Rameez Rahman, Bojan Karlas, and Karl Aberer. 2016. The Curious Case of the PDF Converter that Likes Mozart: Dissecting and Mitigating the Privacy Risk of Personal Cloud Apps. *CoRR* abs/1608.05661 (2016). http://arxiv.org/abs/1608.05661

[5] Ruogu Kang, Laura Dabbish, Nathaniel Fruchter, and Sara Kiesler. 2015. "My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. USENIX Association, Ottawa, 39–52. https://www.usenix.org/conference/soups2015/proceedings/presentation/kang

[6] Predrag Klasnja, Sunny Consolvo, Jaeyeon Jung, Benjamin M. Greenstein, Louis LeGrand, Pauline Powledge, and David Wetherall. 2009. "When I Am on Wi-Fi, I Am Fearless": Privacy Concerns & Practices in Eeryday Wi-Fi Use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1993–2002. https://doi.org/10.1145/1518701.1519004

[7] Michelle L. Mazurek, J. P. Arsenault, Joanna Bresee, Nitin Gupta, Iulia Ion, Christina Johns, Daniel Lee, Yuan Liang, Jenny Olsen, Brandon Salmon, Richard Shay, Kami Vaniea, Lujo Bauer, Lorrie Faith Cranor, Gregory R. Ganger, and Michael K. Reiter. 2010. Access Control for Home Data Sharing: Attitudes, Needs and Practices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 645–654. https://doi.org/10.1145/1753326.1753421

[8] Michelle L. Mazurek, J. P. Arsenault, Joanna Bresee, Nitin Gupta, Iulia Ion, Christina Johns, Daniel Lee, Yuan Liang, Jenny Olsen, Brandon Salmon, Richard Shay, Kami Vaniea, Lujo Bauer, Lorrie Faith Cranor, Gregory R. Ganger, and Michael K. Reiter. 2010. Access Control for Home Data Sharing: Attitudes, Needs and Practices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 645–654. https://doi.org/10.1145/1753326.1753421

[9] Erika Shehan Poole, Marshini Chetty, Rebecca E. Grinter, and W. Keith Edwards. 2008. More Than Meets the Eye: Transforming the User Experience of Home Network Management. In *Proceedings of the 7th ACM conference on Designing interactive systems (DIS '08)*. ACM, ACM, 455 – 464. http://doi.acm.org/10.1145/1394445.1394494

[10] Fahimeh Raja, Kirstie Hawkey, and Konstantin Beznosov. 2009. Revealing Hidden Context: Improving Mental Models of Personal Firewall Users. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. ACM, New York, NY, USA, Article 1, 12 pages. https://doi.org/10.1145/1572532.1572534

[11] Gerald Sauer. 2017 (accessed September 7, 2017). *A MURDER CASE TESTS ALEXA'S DEVOTION TO YOUR PRIVACY.* https://www.wired.com/2017/02/murder-case-tests-alexas-devotion-privacy/

[12] Blase Ur, Jaeyeon Jung, and Stuart Schechter. 2014. Intruders Versus Intrusiveness: Teens' and Parents' Perspectives on Home-entryway Surveillance. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 129–139. https://doi.org/10.1145/2632048.2632107

[13] Rick Wash. 2010. Folk Models of Home Computer Security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10)*. ACM, New York, NY, USA, Article 11, 16 pages. https://doi.org/10.1145/1837110.1837125

[14] Wiktoria Wilkowska, Martina Ziefle, and Simon Himmel. 2015. Perceptions of Personal Privacy in Smart Home Technologies: Do User Assessments Vary Depending on the Research Method?. In *Proceedings of the Third International Conference on Human Aspects of Information Security, Privacy, and Trust - Volume 9190*. Springer-Verlag New York, Inc., New York, NY, USA, 592–603. https://doi.org/10.1007/978-3-319-20376-8_53

[15] Jan Wolf. 2017 (accessed September 6, 2017). *Roomba vacuum maker iRobot betting big on the 'smart' home.* https://www.reuters.com/article/us-irobot-strategy/roomba-vacuum-maker-irobot-betting-big-on-the-smart-home-idUSKBN1A91A5

[16] Rayoung Yang and Mark W. Newman. 2013. Learning from a Learning Thermostat: Lessons for Intelligent Systems for the Home. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 93–102. https://doi.org/10.1145/2493432.2493489

[17] Svetlana Yarosh and Pamela Zave. 2017. Locked or Not?: Mental Models of IoT Feature Interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2993–2997. https://doi.org/10.1145/3025453.3025617