

Safety or Security: What do we Notice?



Submitted To:

Project Advisor: Krishna Kumar VENKATASUBRAMANIAN, WPI Professor *Project*

Co-advisor: Lane HARRISON, WPI Professor

Submitted By:

Devon COLEMAN

Christopher NAVARRO

Jean Marc TOUMA

Date: 3 March, 2017

This report represents the work of WPI undergraduate students submitted to the faculty as evidence of completion of a degree requirement. WPI routinely publishes these reports on its website without editorial or peer review. For more information about the projects program at WPI, please see

<http://www.wpi.edu/academics/ugradstudies/project-learning.html>

Abstract

As technology progresses, ever-increasing amounts of non-traditional devices will be connected to the Internet of Things (IoT). Devices in the IoT are capable of notifying their owners of their status in real time. Because of this, these notifications can carry tremendous import. While there is previous work detailing how humans handle security notifications, none has been done for notifications that deal with physical safety. Using an online study, this project explored this gap in research by examining the difference between the two. Despite most metrics being identical, participants consistently responded more correctly to safety notifications than security, suggesting that there is a higher level of comprehension or understanding present with safety versus computer security.

Contents

1	Executive Summary	5
2	Background	8
2.1	Internet of Things Devices	8
2.2	Human-Computer Interaction in IoT	11
2.3	Security in IoT	11
2.4	Notification Effectiveness	12
2.4.1	Notification Types	13
2.5	Previous Studies on Notifications	13
2.6	Mechanical Turk	13
2.7	Conclusion	14
3	Methodology	15
3.1	Experiment Design	15
3.1.1	Primary Task for Inducing Cognitive Load	16
3.1.2	Secondary Task to Test Notification Response	17
3.1.3	Notification Types	18
3.1.4	Study Flow	19
3.1.5	Experiment Metrics	22
3.2	Hypothesis	25
3.3	Pilot Studies to Validate Methodology	25
4	Amazon Mechanical Turk Study 1	27
4.1	Notification Response Time	27
4.2	Notification Response Correctness	28
4.3	Notification Recall Correctness	28
4.4	Complexity Analysis of Notification Content	29
5	Amazon Mechanical Turk Study 2	31
5.1	Notification Response Time	31
5.2	Notification Response Correctness	32
5.3	Notification Recall Correctness	33
6	Discussion and Future Work	36
6.1	Novelty Effect	36
6.2	Comprehension	36
6.3	Future Work	37
6.4	Conclusion	37
	References	38
A	Module 1: Informed Consent Agreement for Participation	40

B	Module 2: Study Introduction	42
C	Module 3: Training Phase	43
D	Module 4: Game	45
E	Module 5: Follow-up Questionnaire/Survey	46
F	Module 6: Score Presentation	47
G	Module 7: Recall Study	48
H	Module 8: Demographics Survey	49
I	Module 9: Final Comments and Code Distribution	50

List of Figures

1	Nest Thermostat	9
2	Smart Fridge UI	10
3	iKettle Map Plot	12
4	Experiment Layout	16
5	Primary Task	17
6	Example Notification	18
7	Participant Score	21
8	Recall Study	22
9	Average Response Time w/ Confidence Intervals	27
10	Notification Response Correctness	28
11	Recall Correctness	29
12	Average Response Time w/ Confidence Intervals	31
13	Notification Response Correctness	32
14	Education Level vs. Notifications Correct	33
15	Notification Recall Correctness	34
16	Type Correctness vs. Grade Level	35
17	IoT Image	42
18	Example Notification	43
19	Example Primary Task with Next Button	44
20	Primary Task	45
21	Secondary Task	45
22	Participant Score	47
23	Recall Image	48

List of Tables

1	Safety Notifications	20
2	Security Notifications	20

1 Executive Summary

As technology progresses and Internet access begins to be regarded as a necessity rather than a convenience, more and more devices will be connected to a network that is commonly referred to as “The Internet of Things.” This Internet of Things, or IoT, is predicted to have an enormous impact on the future of nearly every sector, including business, healthcare, and science [3]. IoT devices range from refrigerators that notify their owners when the milk is going bad to medical sensors that report all health information back to a central server for more personalized treatment. IoT devices can make people’s lives easier and more convenient; however, they are not without risks.

These new devices are only another source of detritus in an increasingly-cluttered stream; now notifications about the status of the groceries in the fridge will vie for attention with new friend requests and emails. The difference between these two are that in the world of the IoT, these notifications can have real, dangerous consequences. For example, if a notification from a fire-detection system appears but is buried under all the other notifications that often collect during the day, there may well be real property damage because no action was taken. Important notifications are now at risk to concealment because of the additional noise.

This project focused around wearable medical devices, and the risk created by adding these devices to the world of IoT. These devices are rapidly becoming much more complex than the simple devices most are familiar with; in fact, they are beginning to resemble wireless sensor networks and can be used for ubiquitous health monitoring [15].

Imagine a hypothetical patient who has had diabetes all his life. Every time he eats, he has to calculate the amount of insulin required to counteract the new sugar in his bloodstream. In a man without diabetes, the pancreas automatically calculates and releases the correct amount of the hormone. With the advancements in sensor networks, the patient can now attach his insulin pump to a continuous blood glucose monitor, automating the process of counteracting a rise in blood sugar while also providing a convenient, instant way to survey blood glucose readings on his base station. This base station is usually a smartphone or wearable medical device. In a perfect system, this would function flawlessly, always giving the correct amount of insulin; but, what if the sensor malfunctions? What if the system is hacked? Then the pump is taking or being fed erroneous readings and acting on them. Furthermore, no system is perfectly secure and safe from attacks.

In order to accommodate the mobile environment of the body, these sensors are becoming wireless. These wireless body area networks (WBANs) are then much more susceptible to attack [11] as they are no longer on a closed circuit. A malicious actor has many vectors of attack on this system. She can spoof sensor readings, making the patient’s glucose monitor notify the pump that there is an extremely high level of glucose in the blood and causing the pump to release a large amount of insulin. She

can attack his pump directly to achieve the same result, or break into the base station to display incorrect readings and cause the user to manually override the pump to counteract. The result of all of these is potential bodily harm and even death.

This is all made possible due to the wireless nature of these networks [13]. Thankfully, there are ways to detect attacks such as these. Methods are being developed that use complementary biometric signals, such as blood pressure readings in correlation to heart rate, to detect if the data from one sensor is incorrect or malicious in some way [17]. The next question, then, is why can't the responses to these errors be automated? What is stopping computers from handling all this data and selecting the best course of action?

The answer to this question is much simpler than it may seem. In a word, it is the unpredictability of humans and our lives. Software is incapable, without ridiculously expensive and expansive machine learning training, of figuring out what a human can know and plan for intuitively. For example, our diabetes patient may know that he will eat more in the future, and be able to preemptively counteract the corresponding rise in blood sugar resulting in a more stable reading. Current research is working only on predicting blood glucose levels between 30-60 minutes in the future, and this prediction is still only 42% accurate. Thus, a human element is still required.

Since a human element is required, there must be a method of communicating data to this element. In typical wearable or portable devices, information is conveyed through use of a notification. This is true for both medical & Internet of Things devices despite poorly fitting the source and type of data provided by these sources.

While much research has been done into the way humans treat security notifications (such as SSL warnings in many popular browsers [19]) as opposed to merely informational ones, none has been done into how we treat notifications that involve personal safety such as those from a WBAN. This background was necessary in order to fully understand presenting notifications relating to medical devices. In order to more accurately study the effect, the study was slightly altered to source the notifications from IoT devices as the types of notifications they provide are more relatable to the average person. As computer safety can parallel real-world safety, this study serves as a first step into exploring this new class of notifications, by comparing human reactions and recall of both types to see if there is a difference.

In order to research this potential difference, the team needed some method of simulating the presence of multiple IoT devices while also reaching a large, diverse group of participants. Amazon's Mechanical Turk provided the perfect platform for the task. This study was set up in a manner similar to those used to study security notifications, utilizing a 'primary task' to simulate cognitive load while presenting different notifications to the participant that had a single correct answer and two incorrect answers. As previous research has shown that very little attention is paid to these types of notifications [3], an effort was made to spell out the correct response in the notification text to provide participants who paid attention to the notification the right answer.

Two rounds were made of the study, with the second having slightly modified notifications to standardize the grade levels of the texts. In each round, nearly every metric was identical; for example, participants remembered the notifications equally and they answered them in about the same amount of time. The only clear and significant difference was in the correctness of the response. While each had poor statistics in keeping with studies that show most people will just click through notifications such as these, safety notifications were consistently answered correctly 20% more often. Some potential explanations proposed for this are the novelty effect and the differing complexities of notifications.

2 Background

Before explicitly laying out all the methodology of the study, it is necessary to understand some of the technologies and definitions that will be used throughout the paper. This section will serve to introduce the reader to these terms and technologies to allow for easier understanding and analysis of the paper.

2.1 Internet of Things Devices

The first, most important term is the "Internet of Things." This term was likely coined by Kevin Ashton [2] while working for Proctor & Gamble in 1999. He explains the concept simply as humans hooking up another source of information to the Internet. This source of information is the creation and connection of many "smart objects," physical "things" that have some increased capacity to affect or record their environment. These "things" can be anything, from a household refrigerator to each and every lightbulb in a home.

With a network connection and potentially an increased sensing capacity, these "things" can send status updates to their owners on specific criteria. For example, the smoke detectors could notify the homeowner that a fire has been detected and that they should call the fire department to control it. This has incredible potential to improve people's lives; now, instead of waiting until someone notices the fire, the homeowner has the information directly at their fingertips and can immediately take action to rectify the situation. These notifications have yet to be standardized same as the IoT space, but the actionable set of IoT notifications can generally be classified as 'safety.' This is analogous to browser SSL warnings and antivirus alerts being classified as 'security,' but carries a sense of heightened risk for people and belongings instead of just technology.

IoT devices are just getting their start; while they haven't begun heavy growth just yet, some believe that they may be on the cusp of a boom similar to that of the smartphone [5]. This belief places enormous importance on the sector as a whole; with such real-world consequences possible, it is of paramount importance that the IoT and its effects are understood as well as possible before widespread adoption. Part of these studies must involve interaction with IoT devices and how their data is presented and manipulated.

Simple examples of IoT devices include the Nest thermostat, which is a smart thermostat capable of not only reporting and setting temperature remotely but also learning the user's schedule and creating a temperature schedule accordingly, and a Samsung smart refrigerator which allows for notes, weather display, and fine-grain temperature control as well as all the usual applications of a refrigerator. Images of both devices are below.



Figure 1: Nest Thermostat

This figure is an image of the Nest thermostat developed by Nest Labs. It uses its temperature sensors and settings as well as an internet connection to develop an understanding of its user's habits and through that, an automatic temperature schedule. Image found at <https://nest.com/videos/thermostat/meet-nest-thermostat/saves-energy/stills/saves-energy-opening-US-d1172ff47b.jpg>.



Figure 2: Smart Fridge UI

This figure is an image of a UI developed in 2012 by Samsung for one of their first models of smart refrigerator. While this model only supports notes, weather, a calendar, and similar apps, more recent models allow for direct-to-store ordering and food freshness estimation. Image found at <https://shinesg.files.wordpress.com/2015/10/samsung-wifi-smart-fridge-2012-01.png>.

2.2 Human-Computer Interaction in IoT

There has been little research into human-computer interaction as it relates to the Internet of Things; in fact, most research on this topic simply explains how everyone else is doing it [8]. Most interactions with IoT devices or information is done through a traditional graphical interface [12], which provides an experience ill-fitting the type of data provided. Data such as this, which is acquired from a real-world setting, would be better suited for understanding and manipulation through other interfaces such as voice or gesture control. Unfortunately, although related to this project in that we are simulating and presenting IoT data generation for greater control, it is not our focus. Thus, typical interfaces will be used to provide participants with data.

2.3 Security in IoT

As stated before the era of IoT is underway. In fact, it is estimated that by the year 2020 approximately 20 billion IoT devices will be in use [10]. Initially threats of security in the past were mainly concerned with information leakage or loss of service; however, IoT devices have made these threats a lot more non-virtual. Security breaches now have the potential to affect our physical well-being, something that shouldn't be taken lightly.

One example in particular looks at a security flaw recognized within a smart teapot. Although seemingly innocent this device had almost no security whatsoever [16]. A malicious attacker simply had to have a stronger wireless signal than the original network the teapot was connected to. Then, with a simple command the teapot would provide the wireless password for the home's network. As a proof of concept for the attack, an individual went around London plotting the various vulnerable teapots, as seen in Figure 3.

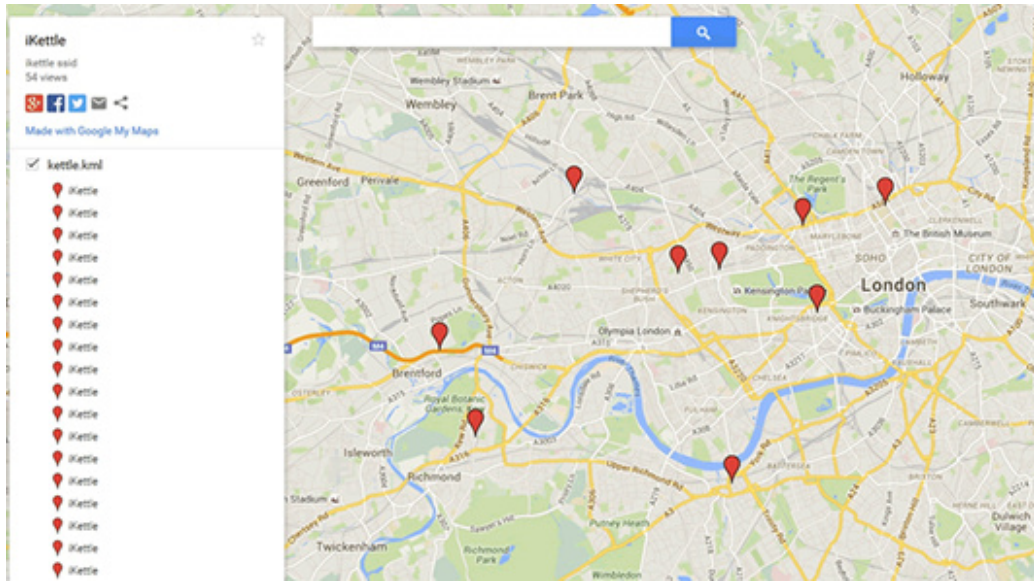


Figure 3: iKettle Map Plot

This figure shows the locations of smart teapots throughout London. These locations were obtained through a security exploit present in the iKettle IoT device.

This example highlights where the level of security in IoT devices currently lie. It can be expected that as more and more devices are created the cracks in security will too. Although, the teapot example was non-life threatening it provided a foothold that attackers could use to pivot to more important devices. With even the smallest possibility of a security breach more measures should be put into place to protect the IoT users. Even with the tightest security, however, attacks can and will happen. When they do, notifications will be provided to the user, so it is still very important that we understand how they will react to them. Our project looks at the potential of these notifications, and their viability.

2.4 Notification Effectiveness

When looking at the effectiveness of security notifications it is easy to see a trend. Often, many notifications or warnings are simply ignored or disregarded. Studies have tried to redesign these notifications to no avail, either by rewording the notification or simply displaying it differently [19]. Unfortunately, the fact remains that people will consistently respond to security notifications incorrectly; the question, then, becomes if the type of data the IoT provides will cause a different response.

IoT devices allow security threats the opportunity to become more than just threats. These devices, and their associated notifications, can have direct, tangible effects on people and the world they live in. Although security notifications have been studied in great detail, little research has been done when looking at notifications that

deal with personal safety. The aim of this project is to see if security notifications are treated the same or differently as these IoT-related ones, termed safety notifications.

2.4.1 Notification Types

As we are studying the difference between safety and security notifications, it becomes necessary to define what exactly they are. For the purposes of this project, a security notification is any notification that is based in computer security. This means that any notification from an antivirus is treated the same as any kind of warning from a browser such as an improper SSL certificate warning. There has been much research suggesting people simply ignore or barely skim such warnings [3] and this project will study whether a similar effect is present for safety notifications, which are notifications that can affect personal safety such as those triggered by fire alarms or vehicle theft.

2.5 Previous Studies on Notifications

Despite little research being done on safety notifications, the existing research on security notifications still proved useful in many areas. One such area was in designing the experiment itself; the idea of introducing a cognitive load to more realistically study reactions to notifications, for example. The primary task/secondary task dichotomy was modeled after that found in [22] study. Further, it was found that most users click through 50% of SSL warnings in 1.7 seconds [7]. A method of reducing this is proposed and analyzed in [6], where the researchers discovered that habituation is a very high component of this ignorance and, through modifying the appearance of the notifications, they can 'greatly increase notification resistance' to this effect. This led to the development of one of the major hypotheses used to explain the effect found in this paper.

2.6 Mechanical Turk

As the IoT industry is still fledgling, however, the team did not have access to a pool of participants who owned many IoT devices. Therefore, in order to study IoT notifications, it became necessary to simulate having IoT devices for a large number of users who did not possess them. This simulation was performed using Mechanical Turk. Mechanical Turk is a service introduced by Amazon that “*gives businesses access to a diverse, on-demand, scalable workforce*” [1]. When an entity, called a Requester, has a large task that needs to be completed but cannot be reliably or effectively done with computing, such as transcription, they can turn to Mechanical Turk (Turk)’s Workers. Requesters provide the task with some compensation attached and the Workers perform the task, receiving said compensation only when the Requester deems the work valid. This incentivizes the Workers to perform the task correctly while providing work that, while lower in quality than that of a dedicated

professional, is cheaper and can actually be controlled through redundancy to mitigate the loss of quality [14]. Despite the fact that Turk's terms and conditions disallow for collection of any potentially identifying personal information or marketing [5], it is seen as a viable way to perform studies that don't hinge on that information because Workers tend to think and act much like any other large sample would [9]. This can be invaluable for researchers looking to get a large sample size for relatively cheap, and made the decision to use Mechanical Turk for the studies simple.

2.7 Conclusion

This section served to lay the groundwork for an experiment testing the difference between safety and security notifications. To summarize, it was necessary to study participant's reactions to safety notifications in an environment without the confusion of medical devices, so the concept of IoT was brought in to provide a believable, realistic source of notifications. Mechanical Turk was used to reach a large, diverse audience of participants to ensure the study covered as much ground as possible.

3 Methodology

In this chapter, we will explain the methods used when building the study and the ideas behind each of the different modules. We will start with an explanation of the primary and secondary tasks, followed by the list of modules present in the study. Then we will explain all metrics captured on a per-module basis, have a discussion of our early pilot studies, and finally touch briefly on the changes made for the Mechanical Turk studies.

3.1 Experiment Design

As stated before, the study was created to help gather information to measure responses to both safety and security notification types. The data collected would be used to help test whether participants responded differently to safety notifications versus security notifications. By utilizing Mechanical Turk, we were able to quickly acquire participants and obtain legitimate results without using physical IoT objects or actual notifications.

Our team decided to use Node.js coupled with javascript to host our experiment's site and run the experiment. By using javascript all notification alert generation could be done on the participant's computer and all collected metrics could be sent to a remote database for future analysis.

Once the technological design was determined, we began to design the actual experiment itself. Since it would not be correct to assume that humans are always primed and ready to receive and respond to notifications during their day, we decided to implement both a primary task and a secondary task in our study, similar to the experiment in [22]. A screenshot of both the primary and secondary tasks can be seen in figures 5 and 6, as well as Appendix D.

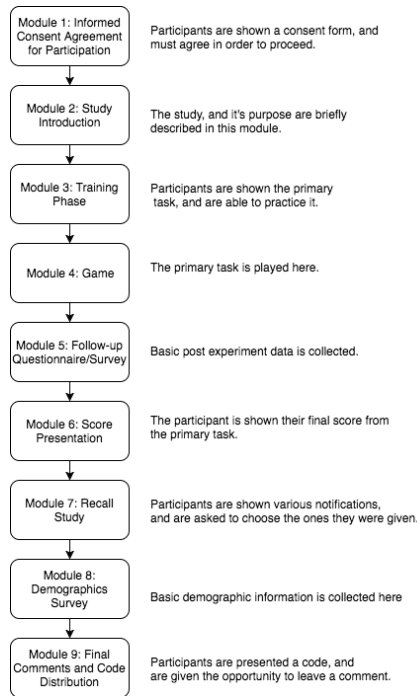


Figure 4: Experiment Layout

This figure outlines the experiment flow. It is broken done into the individual modules.

3.1.1 Primary Task for Inducing Cognitive Load

The primary task in our experiment was meant to simulate the average mental workload a person may experience throughout the day. In order to create this workload our experiment used a word selection primary task (seen in Figure 5). This selection task requires the participant to count the number of times a target word appears in a table. They then must log that number. The primary task awarded or docked points based on a binary scale; a correct answer awarded 100 points and an incorrect answer deducted 100 points. The primary task had two timers: a board timer and the experiment timer. The board timer was started at 45 seconds and was the amount of time the participant had left to count the number of words on the board. This would be reset every time the participant logged a count and would never be greater than the experiment timer. The experiment timer simply showed the amount of time left to perform the primary task.

The use of this mechanism as the study’s primary task is based on previous research related to inducing cognitive stress. This stress was designed to distract participants from the secondary task by simulating everyday workload. Our word selection task is similar to one the the tools designed by [4]. The experiment our design is based off used multiple primary tasks to induce stress. We selected and used a simi-

lar primary task that was seen to be of average difficulty in the experiment.

Experiment Time Remaining - 2 : 04

Primary Task Target Word: Apricot

Board Time Remaining - 0 : 34 **Score: 0**

Avocado	Apple	Asparagus	Aubergine	Avocado
Asparagus	Asparagus	Apple	Arkansas	Asparagus
Asparagus	Apricot	Aubergine	Apple	Asparagus
Aubergine	Apple	Arkansas	Asparagus	Apricot
Asparagus	Arkansas	Arkansas	Arkansas	Arkansas

Figure 5: Primary Task

This figure shows the main primary task that each participant completed. The game involves counting the amount of instances a word appears within the table.

3.1.2 Secondary Task to Test Notification Response

The secondary task in our experiment took the form of predefined safety and security notifications that would appear on screen during the primary task. While the subject was engaged in the primary task, notifications at randomly timed intervals would interrupt the subject. These generated notifications were designed to appear 6 times, showing all 6 notifications, and prevent the subject from interacting with the primary task thus forcing them to respond to the notification. This interruption was also aimed at creating additional stress and anxiety during the experiment, putting more pressure on the user to act [11]. An example of a displayed alert can be seen in Figure 6. The user would then need to respond to the notification by choosing an action from the options provided. The subject would select the option they believed to be the best response to each notification prompt.

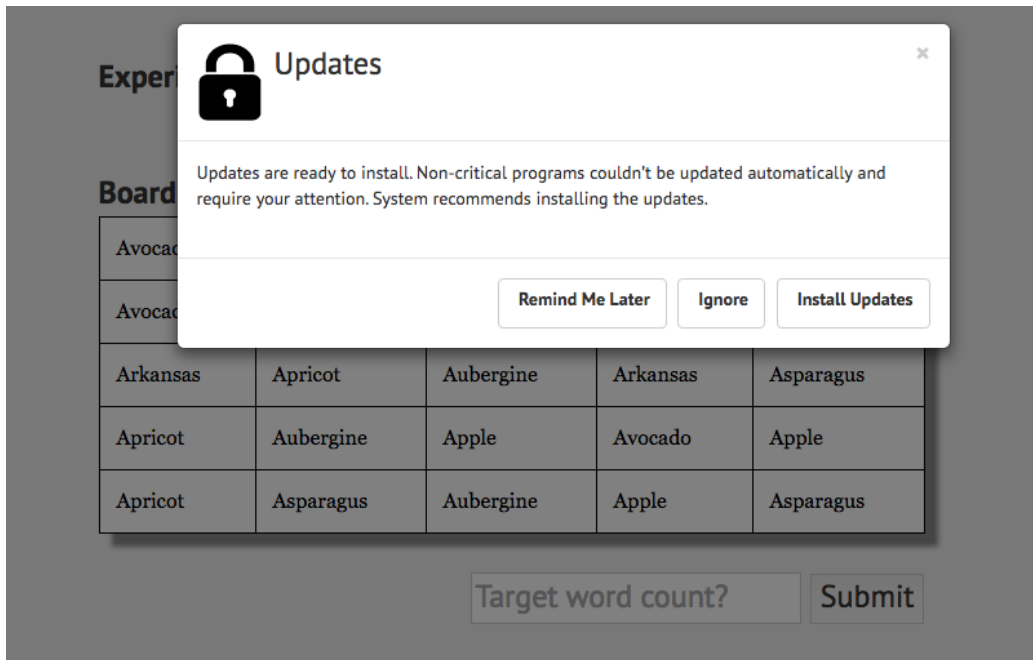


Figure 6: Example Notification

This figure shows the notifications that interrupted the participants during the primary task game. The correct action for each notification was presented within the text. However, to correctly answer the participant was required to read the text.

3.1.3 Notification Types

Since the study focused on the difference between responses to safety and security notifications, it was important to provide real world examples of such notifications to the participants. To achieve this we designed 6 different safety and security notifications that were used in the study. Each type of notification had the same general design; all that differed was the picture in the top right of the notification and the text within. Security notifications were modeled off of those present in operating systems such as Windows and Mac OS and those present in web browsers such as Google Chrome and Mozilla Firefox. There was much less preexisting material to draw from for safety notifications as the industry is still getting off the ground; however there were some examples such as those found in the CarLock protection system [20].

Tables 1 and 2 outline the various notifications that were used in our experiment study. Each table contains a specific notification type; table 1 safety, and table 2 security. Each participant was presented with 6 notifications at random intervals, varying in type. Each notification displayed presented the participant with three actions to take and was designed in a way to hint at the correct course of action. This was done to check whether or not the participant was truly responding to notification or just selecting an answer in order to return to the primary task faster. These choices

were also randomized in the order they appeared to prevent any bias. Finally, a mapping between the perceived severity of each notification was created. In the same way that a disk space warning is less severe than an antivirus warning, a refrigerator going into deep freeze is less severe than a fire alarm. In this way, the standardization between notifications was improved.

3.1.4 Study Flow

Our study on Mechanical Turk is built in such a way that it is broken down into different modules. Each module in our experiment has a different purpose and will be briefly explained in this section (in order of appearance). For more details, such as the exact wording of specific modules, all modules can be viewed in the appendices. There are a total of 9 modules:

- Module 1: Informed Consent Agreement for Participation
 - During this phase the participant provides their Mechanical Turk username and read and agreed to an experiment consent form.
- Module 2: Study Introduction
 - Introduces participants to the concept of the Internet of Things and sets the scenario of the study.
- Module 3: Training Phase
 - Provides exposure to the primary task as well as an example notification. This phase is intended to reduce the learning curve and prepare participants for the actual study. See Appendix C for further detail of the layout of this module.
- Module 4: Game
 - The actual experimentation phase that contains both the primary task and secondary tasks. A screenshot of this phase’s tasks can be seen in Figure 5 and in Appendix D.
- Module 5: Follow-up Questionnaire/Survey
 - At the end of the experiment the participant is asked to complete a small survey, included in Appendix E. This survey focuses on the participant to describe which notification types they saw, and how many of each type they were presented.
- Module 6: Score Presentation

Notification Title	Notification Descriptions	Actions (*correct choice)
Smart Fridge Temperature Warning	Your fridge is extremely cold which may result in supercooling. Raising the temperature will prevent this.	<ul style="list-style-type: none"> • Ignore • Reset Fridge • Raise Temperature*
Fire Alarm Triggered	Smoke particulates have been detected in your kitchen. The system recommends contacting the Fire Department.	<ul style="list-style-type: none"> • Call Fire Department* • Disable System • Ignore
Vehicle Security System Triggered	Your car engine has started without your permission. System recommends engaging the brakes.	<ul style="list-style-type: none"> • Ignore • Lock Brakes* • Call Police

Table 1: Safety Notifications

Notification Title	Notification Descriptions	Actions (*correct choice)
Disk Space Warning	You are running very low on disk space on your primary drive. Disk cleanup suggested.	<ul style="list-style-type: none"> • Perform Disk Cleanup* • Ignore • Wipe Drive
Updates Ready	Updates are ready to install. Some important programs need manual update right away.	<ul style="list-style-type: none"> • Ignore • Remind Me Later • Update*
Antivirus Warning	A file has been discovered to be infected. Infection has been quarantined and file deletion suggested.	<ul style="list-style-type: none"> • Ignore • Disable Antivirus • Delete File*

Table 2: Security Notifications

- The participant’s score on the primary task is presented. A screenshot example can be seen in Figure 7 or Appendix F.

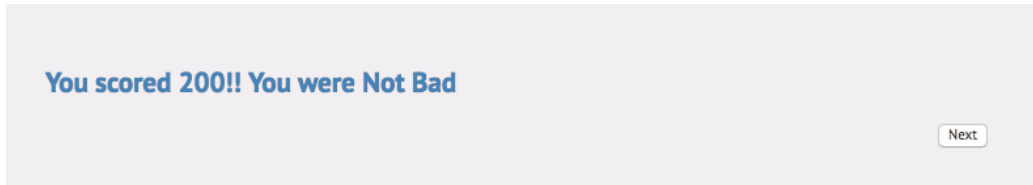


Figure 7: Participant Score

This figure shows the score that participants were shown after the primary task game and survey.

- Module 7: Recall Study
 - The participant is asked a series of recall questions based on the notifications they were shown during the survey. This module it to test whether one notification type is remembered better than the other. See Figure 8 or Appendix G for a screen shot of the recall study.

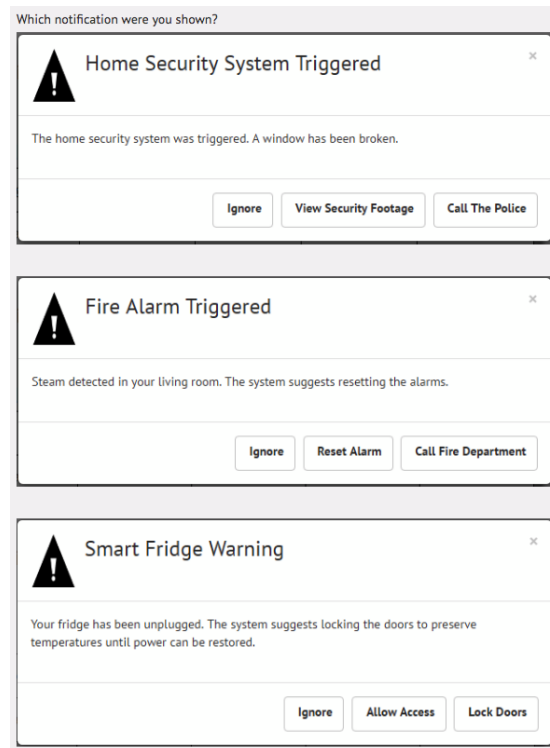


Figure 8: Recall Study

This figure shows the layout and recall notifications that participants were shown during the recall module. This module was intended to determine whether either notification type left more of an impression than the other.

- Module 8: Demographics Survey
 - The participants are presented with a basic demographics survey, included in Appendix H. This data is collected to enable data to be broken down using demographic filters.
- Module 9: Final Comments and Code Distribution
 - The participants are presented with an optional text-box in which they can add any comments or feedback about the experiment. They then press a button to receive their unique ID that they can enter back on Mechanical Turk to receive credit for taking the experiment. A screen shot of this module can be seen in Appendix I.

3.1.5 Experiment Metrics

During the course of the experiment, it was necessary to log various metrics in regards to how participants responded to the various notifications. This section out-

lines all metrics and data collected throughout the experiment as well as giving brief description of each datum, module by module.

- Module 1: Informed Consent Agreement for Participation
 - *Amazon worker ID*: The participant’s Amazon worker ID which will be used to disallow participants from taking this experiment multiple times.
 - *Experiment Start Time*: Timestamp marking the start of the experiment.
- Module 2: Study Introduction
 - No metrics are recorded in this module
- Module 3: Training Phase
 - *Training Module Start Time*: Timestamp marking the start of the training module.
 - *Training Module End Time*: Timestamp marking the end of the training module.
 - *Training Module Total Time(ms)*: Difference between the start and end timestamps recorded for the module.
- Module 4: Game
 - Primary Task:
 - * *Primary Task Start Time*: Timestamp marking the start of the primary task. Starts on every new board generation.
 - * *Answer Provided*: The answer the participant provides for the number of times the target word appears in the board.
 - * *Correct Answer*: The number of times the target word appears in the board.
 - * *Primary Task End Time*: Timestamp marking the end of the primary task. Logged once participant chooses answer for board instance.
 - * *Primary Task Total Time(ms)*: Difference between the start and end timestamps of each board instance.
 - * *Final Score*: Total score achieved by the participant in the primary task.
 - Secondary Task/Notification (Performed approx. 6 times):
 - * *Notification Start Time*: Timestamp marking the start of the notification.
 - * *Type*: Records notification type (safety or security), as well as which specific notification was shown.

- * *Answer Provided*: The answer the participant chose based on the notification.
 - * *Notification End Time*: Timestamp marking the end of the notification.
 - * *Notification Total Time*: Difference between starting and ending timestamps for the notification.
- Module 5: Follow-up Questionnaire/Survey
 - *Notification Types Shown*: Safety, security, or both.
 - *Total Number of Notifications Shown*: The number of notifications the participant recalls seeing in the notification.
 - *Notification Type Shown More*: The type of notification the participant recalls seeing more of, either security or safety.
 - *Number of Security Notifications Shown*: The number of security notifications the participant recalls seeing.
 - *Number of Safety Notifications Shown*: The number of safety notifications the participant recalls seeing.
 - Module 6: Score Presentation
 - No metrics are recorded in this module
 - Module 7: Recall Study
 - *Recall Answer (Performed approx. 6 times)*: Participants must select from 3 notification options and attempt to select the one they were shown during the experiment (see Appendix G). Both the participant's answer and the correct answer is logged.
 - Module 8: Demographics Survey
 - *Age*: The participant's age
 - *Sex*: Male, Female or Other
 - *Country*: Participants select a country from a drop down of all countries
 - *Highest degree obtained*: High School, Bachelors, Masters, PhD or other.
 - Module 9: Final Comments and Code Distribution
 - *Feedback/Comments*: Participants have the option to leave feedback on the experiment. Any comment is logged and recorded.
 - *Code Distributed*: Amazon Mechanical Turk code given for verification of completion of experiment.

- *Experiment End Time*: Time-stamp which marks the end of the experiment.
- *Experiment Total Time(ms)*: Difference between the starting and ending time-stamps.

3.2 Hypothesis

This study was intended to detect if there is a difference between human responses to safety and security notifications. The hypothesis that was being tested and the corresponding null hypothesis, therefore, is outlined below.

H1: There is some statistically significant difference between human responses to safety and security notifications.

H0: There is no statistically significant difference between human responses to safety and security notifications.

If there is a statistically significant difference between the two, the null hypothesis will be rejected. If there is not, the opposite will happen.

3.3 Pilot Studies to Validate Methodology

In order to ensure a smooth study deployment to Mechanical Turk, we opted to host small pilot studies with smaller groups of students. These consisted of two rounds. The first round had five participants while the second round had ten. The purpose of these pilots was simple: acquire some preliminary data and try to discover any faults in the experiment that should be fixed before pushing to a wider participant pool on Turk. The insights this data provided were the following:

- In some instances participants would get stuck in the practice phase for a while before realizing that they hadn't started the experiment yet. In order to fix this issue, the practice phase was made more visually distinct with a very obvious next button and explanation stating that it was a practice module, not the actual study.
- Another observation was confusion on what exactly safety and security notifications meant. Since the study revolved around examining these two types of notifications it would be important that participants understand each. In order to address this, descriptions were provided of each type before the practice module.
- During the first pilot study it became apparent that our initial logging strategy would be difficult to upscale once we started receiving larger amounts of responses. This proved to be a quick fix, but an important one. Each participant was changed to be stored as a separate data object, which contained data

sub-objects on a per module basis. This strategy enabled faster analysis on the data by making it far easier to process and analyze.

- Finally, the data allowed for the calculation of the Turk participants' compensation by using the measured average time that participants took to complete the study. The study took, on average, about 6-8 minutes. This length was used to help determine a fair compensation by dividing it into US minimum wage, producing a payment of \$0.75-0.80 per participant.

4 Amazon Mechanical Turk Study 1

Once all data was collected and all modifications made to the study, it was time to push to a larger audience. The experiment was released on Mechanical Turk, with the trial allowing only a predetermined amount of participants to complete it. Each participant was then compensated for the value calculated above. This trial consisted of 30 participants.

4.1 Notification Response Time

One area in which there appeared to be very little variance between safety and security notifications was in the average response time for each notification. This response time is defined as the time difference between the notification being shown and interacted with. This data is shown in Figure 9.

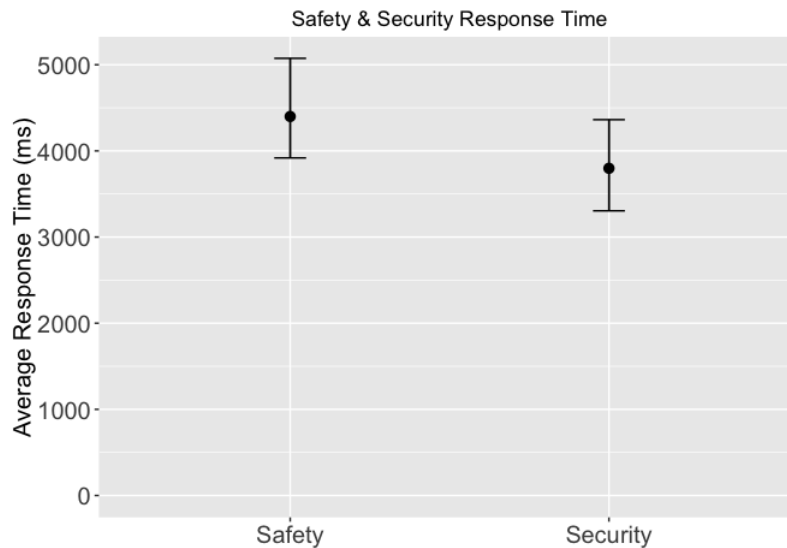


Figure 9: Average Response Time w/ Confidence Intervals

This chart shows the difference in response time between safety and security notifications. This response time was the measure of how long it took participants to select an action per notification.

While safety notifications may appear to have a slightly longer response time, this is not statistically significant.

4.2 Notification Response Correctness

However, not all the data showed this much homogeneity. One area in which there appeared to be a difference was in notification correctness. The data for the first study is shown in Figure 10

There appears to be a significant difference between safety and security. On average, people respond to safety notifications with the correct answer at least 20% more often than they do with security notifications. With all other data being more or less identical between both types of notifications, this difference becomes even more stark.

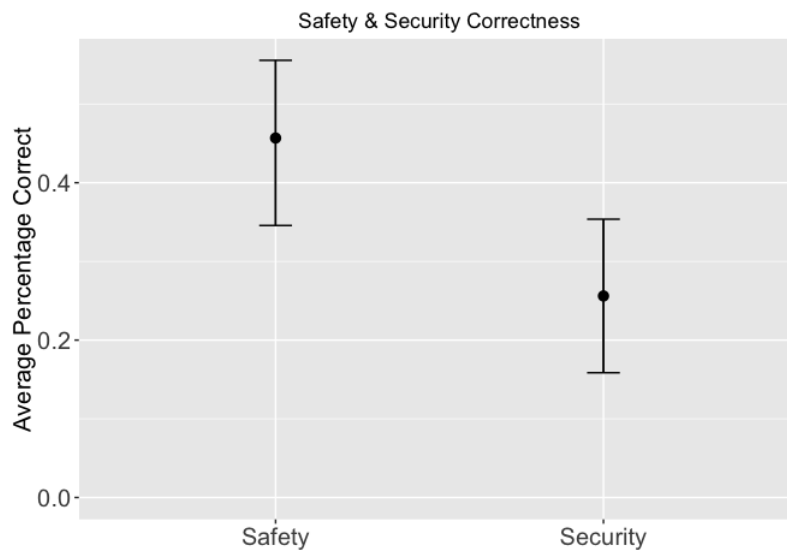


Figure 10: Notification Response Correctness

This chart shows the difference in notification correctness between safety and security notifications. This metric was the measure of how often participants could correctly respond to the notifications they were shown.

4.3 Notification Recall Correctness

In a similar vein, the notifications were recalled with almost identical precision. In Figure 11 there appears to be even less difference between these two confidence intervals than in those generated from the response time.

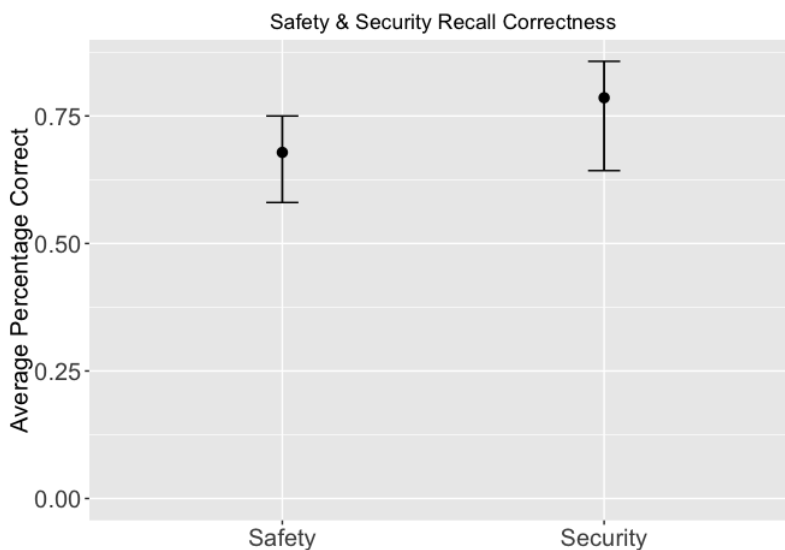


Figure 11: Recall Correctness

This chart shows the difference in recall correctness between safety and security notifications. This metric was the measure of how often participants could correctly recall the notifications they were shown.

4.4 Complexity Analysis of Notification Content

After receiving the results from this study, we wanted to confirm the validity of the results, particularly the response correctness. In order to help remove any room for confounds an additional change was made as a result of this first study. The notifications needed some form of standardization to more directly compare the content of each notification type, rather than just the wording and severity.

Participants tended to respond more accurately to safety notifications by a small margin of about 5-10%. One possible cause proposed for this effect was that the provided safety notifications have less word complexity than security; thus people simply get them correct more often because they have a better understanding. In order to eliminate this potential cause, there was a need to standardize the complexity of safety and security notifications. Thus, each notification was run through a readability analyzer located at [18] to calculate the Dale-Chall readability score. This score is generated by counting the number of so-called ‘complex’ words in the passage and performing a mathematical transformation. Complex words are defined as those not in a list of 3000 that a 4th grader could reasonably be expected to know. This score is one of the most accurate readability metrics [18].

In order to standardize the notifications, a simple formula was created: each security notification should have a readability score that matches the score of one safety notification, with a one-to-one mapping. This way participants are exposed to

a range of notifications from simple to complex, but for each simple safety notification there is also a security notification. This served as an easy method for eliminating any study score differences that may be caused by differences in comprehension. Furthermore, an attempt was made to synchronize the perceived 'severity' of the notifications; in the same way that an update notification would be taken less seriously than an antivirus notification, a fire alarm notification would carry more weight than a refrigerator that is running somewhat cold. With these additional change, our second study could be deployed, and would allow for a higher degree of certainty in the results.

5 Amazon Mechanical Turk Study 2

The experiment was released a second time on Mechanical Turk in order to solidify previous results. A power analysis performed on the data obtained from the first trial determined a test of approximately 100 participants would be needed to detect a Cohen's d value of 0.2. Thus, the team decided on a trial of an additional 60 participants to bring the total up to 100.

Cohen suggested that $d=0.2$ be considered a 'small' effect size

5.1 Notification Response Time

As seen in our first experiment there was a small but finally insignificant difference in response times for the notification types. Figure 12 shows no statistically significant differences between the two, and indeed the gap between them closed completely.

This result helps indicate each notification type took approximately the same time for a participant to respond to. This result helps show that the attempts to make each notification unbiased in regard to length and grade level were successful, as the gap between the response times in the first study has been closed.

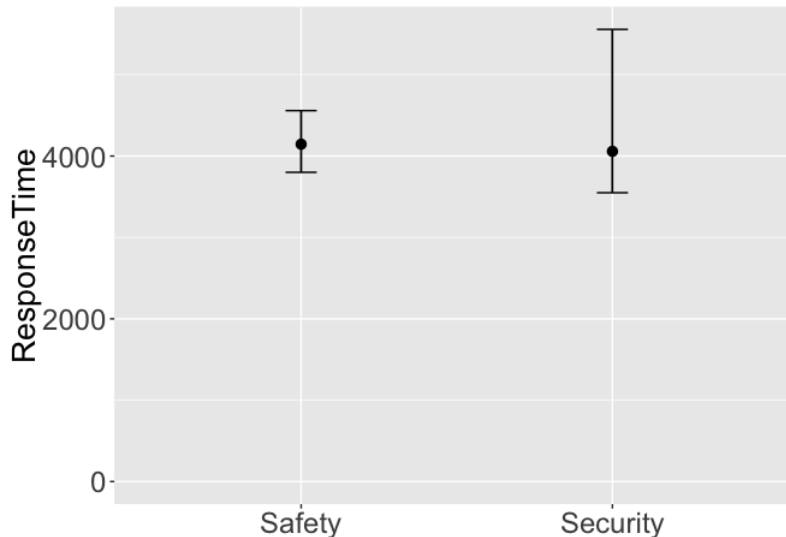


Figure 12: Average Response Time w/ Confidence Intervals

This chart shows the difference in response time between safety and security notifications. This response time was the measure of how long it took participants to select an action per notification.

5.2 Notification Response Correctness

As seen before in the previous study, there was a significant difference in response correctness for each notification type. This result, shown in Figure 13, mimics our first experiment with a greater effect; however, with the additional grade level data it was possible to step into and analyze this result further.

First, we specifically looked at each participant's educational level (seen in Figure 14). We compared each education level with the average notifications they answered correctly. No education level significantly outperformed the others, and the difference in notification correctness was still present in all levels.

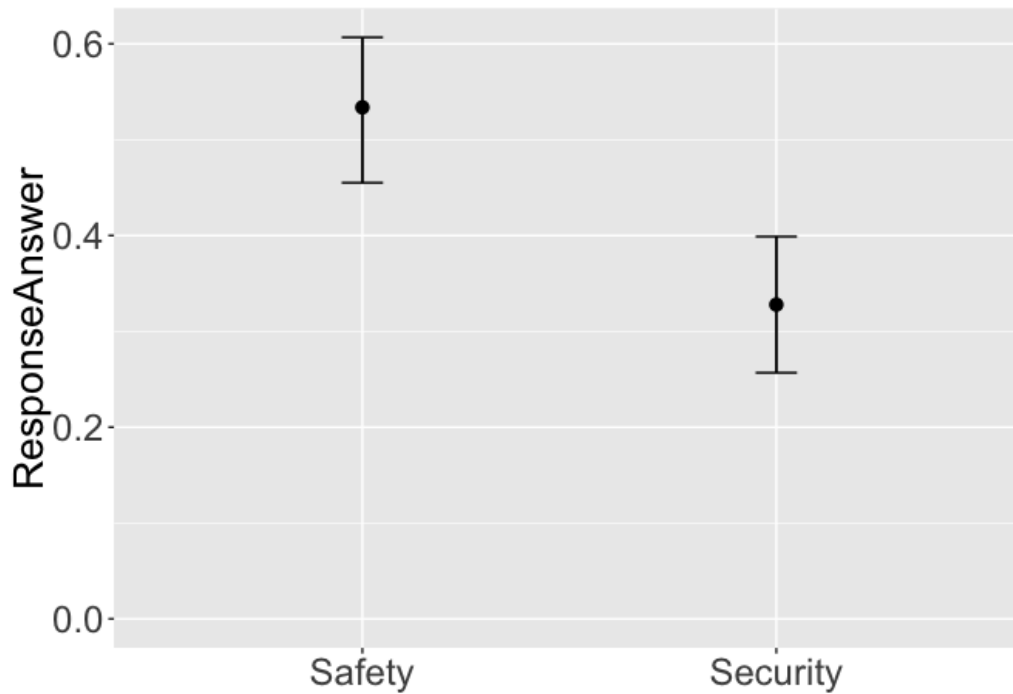


Figure 13: Notification Response Correctness

This chart shows the difference in notification correctness between safety and security notifications. This metric was the measure of how often participants could correctly respond to the notifications they were shown.

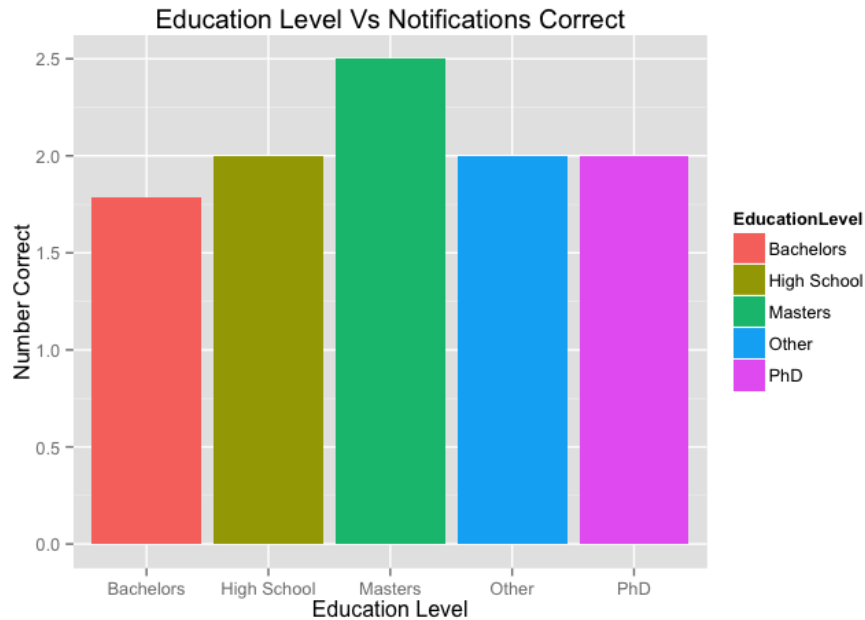


Figure 14: Education Level vs. Notifications Correct
This chart directly compares how correctly participants answered notifications to their education level.

Using the complexity analysis added from the earlier study, we used the grade level, or readability, of each notification for each type. The notifications were designed such that each one had a corresponding notification of the opposite type that matched its grade level score. This was done in order to reduce the possibility that the safety notifications were merely simpler than the security notifications resulting in more participants getting them correct. Figure 16 shows the correctness for each notification type.

5.3 Notification Recall Correctness

Similar to notification response times, recall correctness for each type was almost identical. Figure 15 shows that each type had almost no difference whatsoever during the recall portion of the experiment.

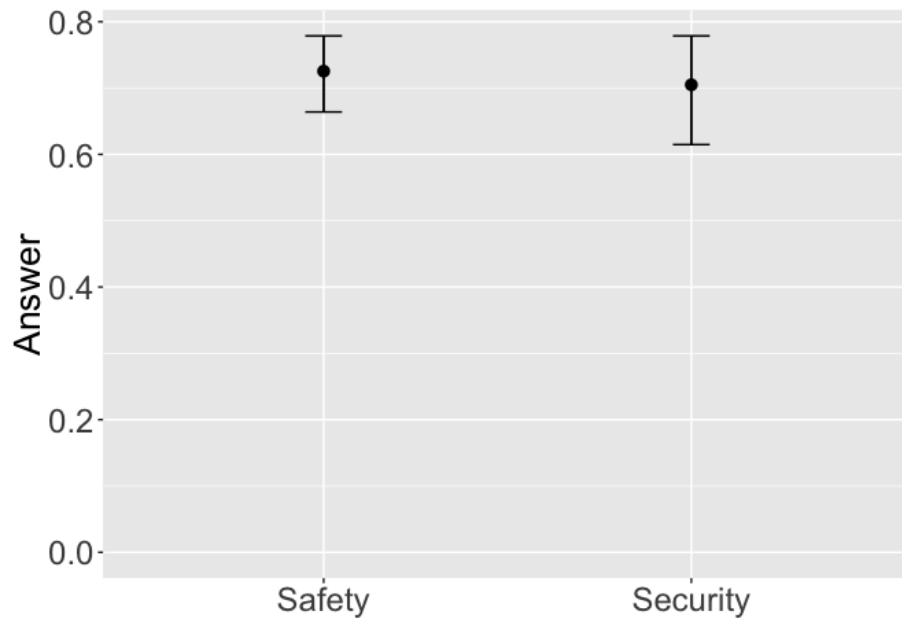


Figure 15: Notification Recall Correctness

This chart shows the difference in recall correctness between safety and security notifications. This metric was the measure of how often participants could correctly recall the notifications they were shown.

The next step was to then analyze the grade levels of all notifications. Using the complexity analysis described in earlier sections, we determined the grade level, or readability, of each notification for each type. Then, notifications were designed such that each one had a corresponding notification of the opposite type that matched its grade level score. This was done in order to reduce the possibility that the safety notifications were merely simpler than the security notifications resulting in more participants getting them correct. Figure 16 shows the correctness for each notification type reflects and even shows a greater effect than our initial finding.

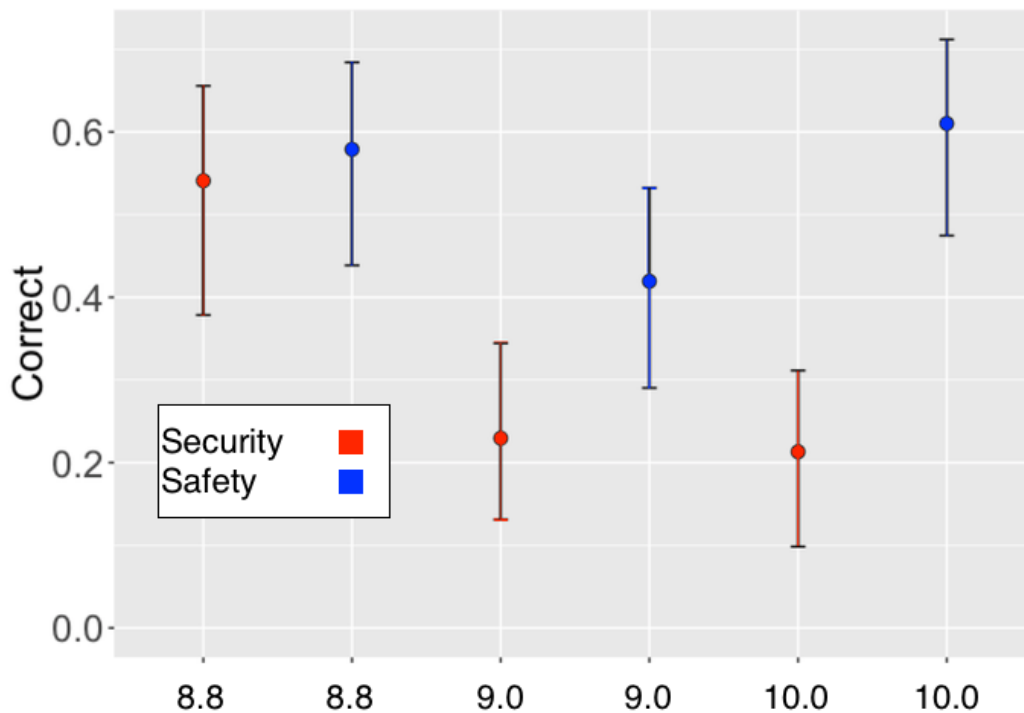


Figure 16: Type Correctness vs. Grade Level

This confidence interval compares safety and security notifications correctness based on their grade level score for word complexity (x-axis). As grade level increases the confidence intervals overlap less and less.

Interestingly enough as grade level increases the difference between safety and security becomes more pronounced. This suggests that participants have an easier time understanding safety notifications regardless of complexity. Some more analysis was performed, searching for other potential ways the difference could be caused; areas such as age affecting correctness, primary task scores affecting the same, and a few others were explored and no trend was found with any of them.

6 Discussion and Future Work

As technology advances, more devices will become part of The Internet of Things (IoT). With that in mind, notifications from these new devices may need to be taken more seriously. This is especially true when these notifications pertain to the physical safety of a person or a person's belongings. Traditional research has been done on how security notifications are handled; there is still, however, a gap in this field when looking at safety notifications.

This study explored the potential difference between the two types of these notifications: safety and security. Through a series of small pilot studies, and two larger studies through Amazon Mechanical Turk our experiment was able to be refined, and gather useful data. The study's participants were shown and responsible for responding to notifications of both types, while being subjected to a primary task to induce cognitive stress. The data collected from this experiment showed one key difference between the two notification types examined.

Practically all data metrics collected, except correctness, were equivalent between security and safety notifications. With this finding it is safe to say there is some effect at work here. Even with attempts to control for potential complexity differences, the difference in correctness still remained. This left only a few possible explanations for the effect. The two believed most likely were the novelty and comprehension effect.

6.1 Novelty Effect

One possible reason why there is a significant difference in response correctness for safety notifications is the novelty effect. Participants are used to seeing security notifications, and in a high percentage of cases, dismissing them [6]. In the cited study, it was discovered that the novelty effect could be reduced drastically by repeatedly altering the notifications, providing a new appearance each time. In a similar manner, the novelty of never having seen a safety notification before may play a role in participants not dismissing them out of hand, thus paying them more attention and answering them correctly.

6.2 Comprehension

Another possible reason why there is a difference in response correctness could be that, all things being equal, participants simply comprehend safety notifications easier because they deal with subjects that are less esoteric to the common user. 'Your car has been broken into' is far more relatable than 'Your computer has been broken into' because the common user is more exposed to cars and how breaking into one might work. However, this is extremely difficult to empirically measure and would have required at the very least launching a third, larger study.

6.3 Future Work

There are a few areas where future work would be a benefit in understanding this effect, as not every potential confound was perfectly controlled for. This would allow isolation of the effect if it does exist.

One way to control for the novelty effect seen in our analysis is to have the same participant(s) take the experiment multiple times as seen in [21]. This should reduce the novelty effect and, if it is a factor, the difference in notification correctness. Mechanical Turk is not set up in such a way as to allow requesters to force completion of a previous HIT, so the team did not attempt to do so.

In order to reduce effects caused only by differing complexities, efforts can be made in future studies to improve the homogeneity of the notifications that are presented. This will force comparisons only between the content of the notification types and not the complexity, hopefully providing a more pure effect if it exists.

6.4 Conclusion

This project's goal was to identify and a research a gap that dealt with discovering how humans respond to different notifications. These notification types pertained to both personal safety and computer security. In order to fill this gap, a study was developed on Mechanical Turk that simulated notifications from IoT devices and recorded participant reactions. With this data, a statistically significant difference in notification response correctness was discovered: on average, participants responded approximately 20% more accurately to safety notifications. All other metrics were equal, suggesting a difference in comprehension that could be explained by either the novelty effect or differences in understanding. Our contributions are simple: we have begun to expand our understanding of human responses to safety notifications, a critical concept in the future of the IoT as these devices become more common.

References

- [1] Amazon. Amazon mechanical turk requester. https://requester.mturk.com/tour/how_it_works, 2016. Online, accessed April 2016.
- [2] Kevin Ashton. In the real world, things matter more than ideas. <http://www.rfidjournal.com/articles/view?4986>, 2009.
- [3] Telecommunications Industry Association. Realizing the potential of the internet of things: Recommendations to policy makers. https://www.tiaonline.org/sites/default/files/pages/TIA-White-Paper-Realizing_the_Potential_of_the_Internet_of_Things.pdf, 2015.
- [4] Brian P. Bailey and Joseph A. Konstan. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state, 2006. Attention aware systems Special issue: Attention aware systems.
- [5] Harald Bauer, Mark Patel, and Jan Veira. The internet of things: Sizing up the opportunity, December 2014.
- [6] Bonnie Brinton Anderson and C. Brock Kirwan and Jeffrey L. Jenkins and David Eargle and Seth Howard and Anthony Vance. How polymorphic warnings reduce habituation in the brain - insights from an fmri study.
- [7] Devdatta Akhawe, Adrienne Porter Felt. Alice in warningland: A large-scale field study of browser security warning effectiveness.
- [8] Emanuel von Zezschwitz, Daniel Buschek, Axel Hösl, Henri Palleis Hanna Schneider, Tobias Stockinger, Simon Stusak, Sarah Tausch, Andreas Butz and Heinrich Hussmann. Human computer interaction in the internet of things era. https://www.medien.ifilmu.de/pubdb/publications/pub/ZezschwitzHS_2015HCIinIOT/ZezschwitzHS_2015HCIinIOT.pdf, 2015. Online, accessed October 2016.
- [9] Horton, John J., David G. Rand, and Richard J. Zeckhauser. The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14(3):399–425, 2011.
- [10] Yong Ho Hwang. Iot security and privacy: Threats and challenges, 2015.
- [11] Jamil Y. Khan and Mehmet R. Yuce. Wireless body area network (wban) for medical applications, 2010.
- [12] Markku Turunen, Daniel Sonntag, Klaus-Peter Engelbrecht, Thomas Olsson, Dirk Schnelle-Walka, Andres Lucero. Interaction and humans in internet of things.

- [13] Niraj K. Jha Meng Zhang, Anand Raghunathan. Trustworthiness of medical devices and body area networks.
- [14] Rushanan Michael, Aviel D. Rubin, Denis Foo Kune, and Colleen M. Swanson. Sok: Security and privacy in implantable medical devices and body area networks.
- [15] et al. Otto, Chris. System architecture of a wireless body area sensor network for ubiquitous health monitoring.
- [16] Darren Pauli. Connected kettles boil over, spill wi-fi passwords over london.
- [17] Shahnaz Saleem, Sana Ullah, and Kyung Sup Kwak. A study of iee 802.15.4 security framework for wireless body area networks, 2011.
- [18] sarah@datayze.com. Readability analyzer. <https://datayze.com/readability-analyzer.php>, 2017.
- [19] Joshua Sunshine, Serge Egelman, Hazim Almuhiemedi, Neha Atri, and Lorie Faith Cranor. Crying wolf: An empirical study of ssl warning effectiveness: Conducting experiments in a real labor market, 2009.
- [20] Protectus Technologies. 24/7 car protection. <https://www.carlock.co/features-cloud/>, 2016. Online, accessed April 2016.
- [21] Martijn H. Vastenburger, David V. Keyson, and Huib de Ridder. Considerate home notification systems: a field study of acceptability of notifications in the home, 2007.
- [22] Christopher Wickens. Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources, 1983.

A Module 1: Informed Consent Agreement for Participation

Investigator: TBD

Contact Information: TBD

Title of Research Experiment: Safety or Security? What do we notice?

Sponsor: N/A

Introduction

You are being asked to participate in a research study. Before you agree, however, you must be fully informed about the purpose of the study, the procedures to be followed, and any benefits, risks or discomfort that you may experience as a result of your participation. This form presents information about the study so that you may make a fully informed decision regarding your participation.

Purpose of the study

As technology progresses, more and more devices are connected to the internet. With the rise of the Internet of Things (IoT), notifications have become more diverse. However, how people react and respond to notifications of different nature is still not understood. This experiment is a first step towards better understanding, and then adapting notifications to allow for better and faster responses, specifically notifications that relate to a person's health and safety.

Procedures to be followed

You will be given an interactive task to complete that you will be scored on. During this interactive task you must also respond to notifications that will appear throughout the experiment. After a set amount of time the experiment will end and we'll ask you some questions to help provide feedback for the experiment.

Risks to study participants

Given the anonymous nature of Amazon's Mechanical Turk and the common nature of the task, there are no anticipated risks.

Benefits to research participants and others

The possible benefits include helping to discover optimal ways to display information in notification-based alerts.

Record keeping and confidentiality

Records of your participation in this study will be held confidential so far as permitted by law. However, the study investigators, the sponsor or its designee and, under certain circumstances, the Worcester Polytechnic Institute Institutional Review Board (WPI IRB) will be able to inspect and have access to this data. Any publication or presentation of the data will not identify you by your Amazon ID.

Cost/Payment:

This task task approximated 6 minutes to complete. Participants will be reimbursed \$0.80.

For more information about this research or about the rights of research participants, or in case of research-related injury, contact:

Investigator (contact info at the top of this page). In addition, you may contact the IRB Chair (Professor Kent Rissmiller, Tel. 508-831-5019, Email: kjr@wpi.edu) and the University Compliance Officer (Jon Bartelson, Tel. 508-831-5725, Email: jonb@wpi.edu).

Your participation in this research is voluntary

Your refusal to participate will not result in any penalty to you or any loss of benefits to which you may otherwise be entitled. You may decide to stop participating in the research at any time without penalty or loss of other benefits. The project investigators retain the right to cancel or postpone the experimental procedures at any time they see fit.

By clicking below, you acknowledge that you have been informed about and consent to be a participant in the study described above. Make sure that your questions are answered to your satisfaction before signing. You are entitled to retain a copy of this consent agreement.

(Button To Proceed)

B Module 2: Study Introduction

Welcome to our experiment!

In the new world of the Internet of Things (IoT), everything from your car to your refrigerator has a connection to the internet and can be controlled by many of your favorite internet-connected devices.

In this study, you will be provided with a word-counting task to complete. This involves counting the number of times a particular word appears in a table. This will be further explained on the following page.

For the purposes of this experiment, please imagine you have many IoT devices in your home including a home security system, toaster, fridge, and fire alarm system. You will receive notifications within your browser giving you alerts relating to these devices as well as the computing device you are taking the study on.

Please respond to these notifications as though they were real.



Figure 17: IoT Image

(Button To Start Experiment)

C Module 3: Training Phase

Training Phase

In this study, you will be provided with a primary task to complete. It is a fairly simple counting game. You will be presented with a table of words and asked to count how many times a certain word appears in the table. If you are correct, your score goes up. If you are incorrect or run out of time, your score will go down. There is a limit of 2 minutes on the entire test and a limit ranging from 45 to 20 seconds on each individual table.

Below is a sample of the game. Please play until you feel comfortable with how the game operates.

As you play the word-counting game various notifications will interrupt you. It is up to you to respond to these how you see fit.

Click the button below for an example notification.

(Button to show example notification) (Button to proceed to next module)

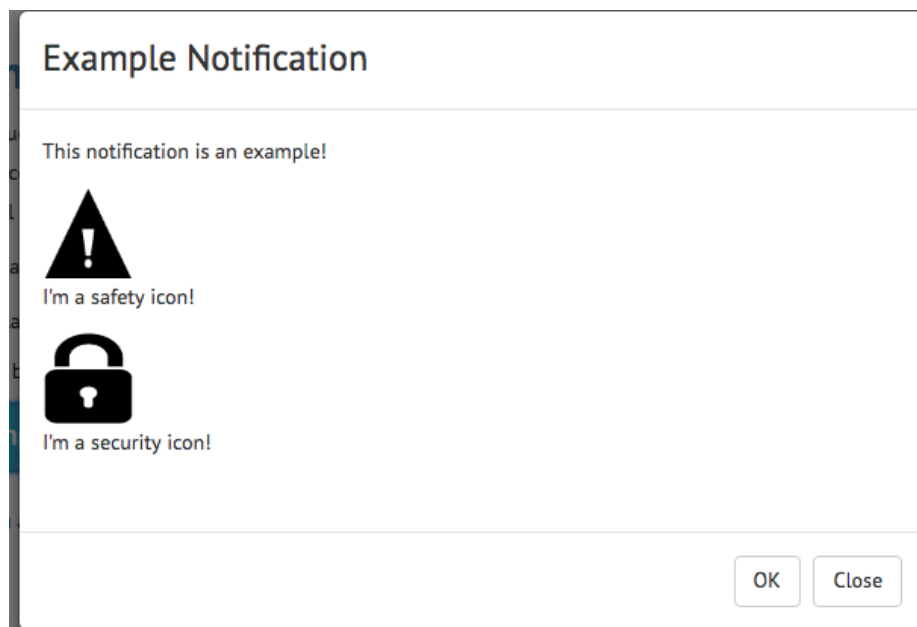


Figure 18: Example Notification

Primary Task Target Word: Apricot

Arkansas	Arkansas	Arkansas	Apple	Apple
Arkansas	Avocado	Aubergine	Asparagus	Arkansas
Apple	Arkansas	Asparagus	Apricot	Apple
Asparagus	Aubergine	Aubergine	Asparagus	Avocado
Arkansas	Arkansas	Asparagus	Arkansas	Asparagus

Figure 19: Example Primary Task with Next Button

D Module 4: Game

Experiment Time Remaining - 2 : 04


Primary Task Target Word: **Apricot**

Board Time Remaining - 0 : 34 **Score: 0**

Avocado	Apple	Asparagus	Aubergine	Avocado
Asparagus	Asparagus	Apple	Arkansas	Asparagus
Asparagus	Apricot	Aubergine	Apple	Asparagus
Aubergine	Apple	Arkansas	Asparagus	Apricot
Asparagus	Arkansas	Arkansas	Arkansas	Arkansas

Figure 20: Primary Task

Exper Board

 **Updates** ×

Updates are ready to install. Non-critical programs couldn't be updated automatically and require your attention. System recommends installing the updates.

Avocado	Avocado	Arkansas	Apricot	Aubergine	Arkansas	Asparagus
Avocado	Arkansas	Apricot	Aubergine	Apple	Avocado	Apple
Arkansas	Apricot	Asparagus	Aubergine	Apple	Asparagus	Asparagus

Figure 21: Secondary Task

E Module 5: Follow-up Questionnaire/Survey

Thank you! The experiment is almost complete. Please fill out the following feedback survey.

We define Security Notifications to be notifications that relate to computer security, and Safety Notifications to be notifications that relate to material and physical safety.

What kinds of notifications were you shown? (Select all that apply.)

- Security
- Safety

How many notifications total were displayed?

(A number box that allows the user to enter any positive number between 0 and 30)

Which notification was displayed more?

- Security
- Safety

How many SECURITY notifications total were displayed?

(A number box that allows the user to enter any positive number between 0 and 30)

How many SAFETY notifications total were displayed?

(A number box that allows the user to enter any positive number between 0 and 30)

F Module 6: Score Presentation

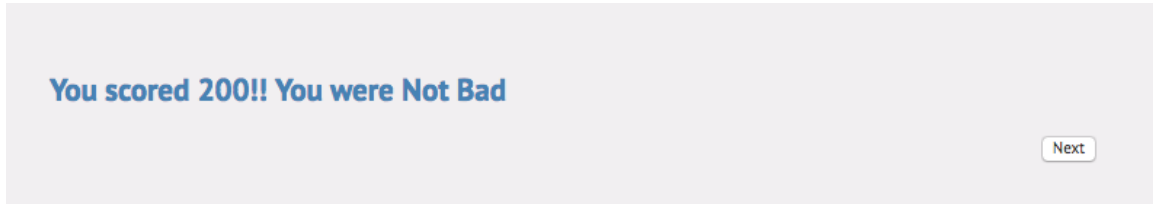


Figure 22: Participant Score

G Module 7: Recall Study

Please try to remember which of the notifications below you were shown. Make your choice by clicking on its picture.

Which notification were you shown?

(Button to show sets of recall notification images)

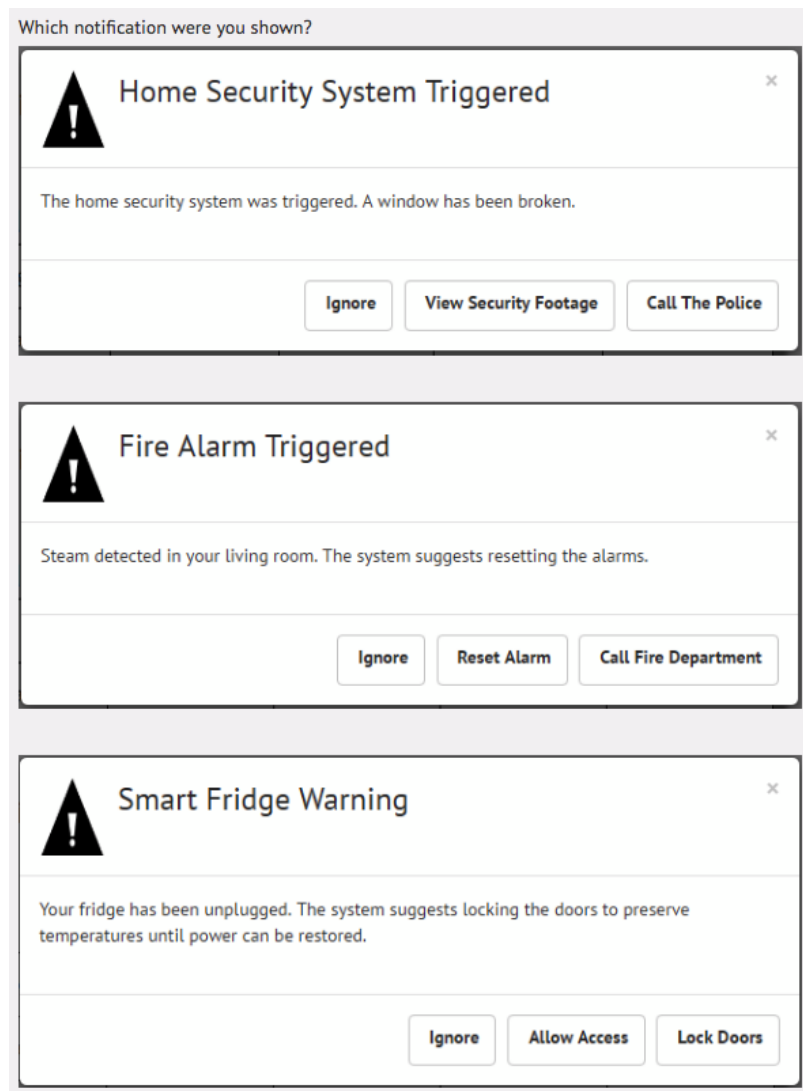


Figure 23: Recall Image

H Module 8: Demographics Survey

Thank you! The experiment is almost complete. Please fill out the following demographics form.

Your age:

(A number box ranging from 1 to 100)

Your gender:

- Male
- Female
- Unspecified

Your country:

(Drop-down list of various countries)

Highest degree obtained (obtained, not pursuing):

- High School
- Bachelors
- Masters
- PhD
- Other

(Button to proceed to next module)

I Module 9: Final Comments and Code Distribution

Thank you again for your participation. Feel free to submit any additional comments below. Click the button to get your code!

(Textbox entry for optional feedback comments)

(Button to submit feedback/show code)

Please copy and paste the following code back on mechanical Turk before closing this window:

EXMPLCODE