The Effects of Multisensory Notifications on User Reactivity

Nicholas A. Croce

March 16, 2017

Abstract

In the past decade, notifications have become an integral aspect of mobile applications. Mobile notifications are used to present users with a variety of information, such as current events, a new message from a friend, a social media comment, an upcoming scheduled event or a scheduled reminder, and much more. Depending on the notification's importance, the mobile user's current context, or the fashion in which the notification is displayed, users can react immediately, slowly, or not at all to a notification. Typically, mobile application developers are limited by the capabilities of a mobile device when notifying their users (the device's sound output, text banners, etc). Developers of traditional alarm clock applications are even more limited, as they rely strictly on the mobile device's sound output to wake sleeping users with their notifications.

Here, I propose an alarm clock application that alerts multiple user senses, and investigate the extent to which these ”multi-sensory notifications” affect user reactivity. To do so, I integrate third party software into my application, to incorporate both sound *and* light in my alarm notifications. I believe that presenting users with a multi-sensory alarm notification will increase their reactivity, when compared to their reactivity to a typical, uni-sensory alarm notification.

In order to test this hypothesis, I conducted a usability study, in which participants woke to both uni-sensory and multi-sensory notifications. Through results I collected by tracking participants' reaction time to the different notifications, I was able to determine whether or not multi-sensory alarm notifications can increase user reactivity.

Contents

1 Introduction

Smartphone and mobile application usage has dominated the first world in the 21st century, and continues to expand its reach as time passes. Over the past 7 years, the number of mobile application downloads per year has increased from 2.5 billion in 2009, to approximately 224.8 billion in 2016 [**10**]. This staggering rate of change comes as no surprise from an economic perspective, because demand for mobile applications continues to heighten as individuals become more and more reliant on smartphones and mobile applications in each coming year.

The roaring increase in the development and usage of mobile applications has created an ever-increasing interest in Human-Computer Interaction (HCI) research [**1**]. Prior to the mobile movement, a vast amount of HCI research was centered around frameworks for desktop-based usability studies. However, the mobile movement created a demand for change and flexibility in the HCI research community. As a result, HCI research has been making a gradual shift towards mobile Human Computer Interaction (mHCI) research, providing somewhat of a framework for usability evaluations of mobile applications [**5**] [**12**].

Because notifications have become an integral aspect of mobile applications, notification system usage and usability has attracted a significant amount of focus from HCI researchers in recent years [**8**] [**6**]. However, there is still room to grow within this field of study. Most of the existing notification-based research focuses on the extent to which a notification system interrupts a user's primary computing task [**2**] [**9**]. Here, I am particularly interested in alarm clock notifications, which are used to wake sleeping users. With this use case, there is no need to assess the interruption effects on users' primary computing tasks, because users are asleep when the notification is presented.

When creating notifications, mobile application developers are typically limited to the capabilities of a mobile device, such as the device's sound output, text banners provided by the OS, etc. When this is the case, typical alarm clock applications must rely on sound output as their only method for waking their sleeping users. What I want to study is the extent to which incorporating a third party developer, in order to add light to an alarm notification, can affect a user's alarm clock experience. Does exposure to light prior to waking heighten user reactivity? Can this multi-sensory notification reduce morning-time grogginess? Are users more or less comfortable being woken up with a multi-sensory notification? To provide answers to these questions, I will conduct a usability study using two applications I have developed, *SmartAlarm* and *BasicAlarm* (see Appendix Sections A and B), to study the effects that multi-sensory notifications have on user reactivity. Participants will be asked to spend 2 days waking with the *BasicAlarm* application, and two days waking with the *SmartAlarm* application. In the morning, they will be asked to fill out email-based surveys that I administer to them, to assess their experience with each type of notification. By comparing the effects that *SmartAlarm's* notifications have on user reactivity with the effects that *BasicAlarm's* notifications have on user reactivity, I should be able to determine whether or not a multi-sensory alarm notification can actually increase user reactivity.

2 Background and Related Work

There hasn't been much HCI research conducted on multi-sensory notification systems with respects to mobile applications. However, as stated earlier, there has been an extensive amount of HCI-based research on usability studies and effective notification systems in the past decade. In addition, a vast amount of psychological and biological studies have been performed regarding sleep patterns, as well as the effects of pre-awakening light exposure on heavy sleepers. In this section, we will cover relevant research and studies that pertain to this thesis.

To begin, we discuss existing research on mobile application usability studies. Alshehri and Freeman [**1**] address user-based evaluation of mobile application usability as a relatively new concept for HCI research. They explain that as computer use shifts to personal empowerment, usability evaluation must become more flexible. It is explained that while laboratory studies are better for collecting high-quality data regarding the usability of the application, field studies are better for gathering data regarding user-satisfaction; the choice of which type of study to use depends on the research objectives. Fayez and Freeman conclude that in order to get the most accurate assessment of an application's usability, a functional prototype is required, and a field study must be performed to present users with real-life context of use. Zhang and Adipat [**12**] delve further into usability testing of mobile devices. In particular, they go into significant detail regarding the selection of research methodology. Laboratory experiments are more effective for applications that don't require network connectivity. Field studies are more appropriate for studying *user behavior* and *attitude towards mobile applications*. It is concluded that field studies should be conducted in order to enable users to provide feedback on experience with the mobile application in a real-world setting. Additionally, Zhang concludes that the greatest challenge of examining usability of mobile applications is the unreliability of wireless network connections.

In addition to usability testing research, it is necessary to observe existing HCI research on notification systems. McCrickard et al. [**8**] provide an introduction to HCI-based research on notification systems. McCrickard asserts that mobile application developers and designers need a greater understanding of how to deliver efficient and effective notifications. It is concluded that the primary challenge in designing notification systems lays within the limitations of human attentiveness, and that testing should be motivated by the cost of error within a given system. LeeTiernan et al. [**6**] study user reactions to notification systems that have a range of reliabilities. Mobile applications with notification systems face user interface design challenges, because there is limited research to provide developers and designers with a framework for creating notification systems. LeeTiernan et al. find that in creating high-quality notification systems, it is very important that users initially perceive the notification system as reliable. Once a user loses trust in a notification-based system, the user may never regain this trust.

Psychological research on the sleep habits of college students is limited. However, Walter C. Buboltz et al. [**3**] provide research on this subject matter in a study performed on a pool of 191 college students. An eight item approach, commonly referred to as the Sleep Quality Index (SQI), was used to measure sleep habits, as well as an open-ended sleep patterns survey. The SQI takes eight sleep pattern-based factors into account, and allows participants to select one of three options for each factor. The factors taking into account by the SQI, and the three options given for each factor are presented in Table 1.

The additional survey distributed to participants recorded several metrics that are not taken into account by the SQI. This survey took both weekdays and weekends into account, and collected typical bedtimes, wake-up times, hours of sleep received, amount of time needed to fall asleep, etc. Buboltz concluded that most college students meet the criteria for poor sleepers. The most common sleep difficulties reported by college students were consistent morning tiredness, and a consistently difficult time falling asleep [**3**].

With the rapid technological advancements of the 21st century, there is room to grow for improving sleep

SQI Factor	**Metric 1**	**Metric 2**	**Metric 3**
Time to fall asleep (mins)	< 10	$11 - 30$	> 30
Suffered from insomnia	Not past 3 months	$<$ 3 times per week	$\geq$ 3 times per week
Difficulties falling asleep	Not past 3 months	$<$ 3 times per week	$\geq$ 3 times per week
Disturbed nights sleep	Not past 3 months	$<$ 3 times per week	$\geq$ 3 times per week
Waking up during the night	$<$ 1 times per month	$<$ 3 times per week	Most nights
Morning tiredness	Mostly alert	Cannot say	Mostly tired
Waking too early	Not past 3 months	$<$ 3 times per week	$\geq$ 3 times per week
Sleep medicines	Not past 3 months	Occasionally	At least 1 time per week

Table 1: SQI Factors and Metrics. Table information taken from [**buboltz-college-sleep**]

patterns and behavior. Choe et al. [**4**] investigate the opportunities computing presents to support healthy sleep patterns. In a study, Choe found that many participants have goals of improving their sleeping habits. These goals mainly included improving sleep-cycle consistency, as well as reducing grogginess, in order to become a morning person. In their findings, Choe et al. discovered that participants wanted several features to be implemented into future sleep-based technology. One functionality that was perceived as useful to participants is the use of *daylight simulation*, which pertains directly to this thesis.

An extensive amount of psychological and biological research has been conducted regarding daylight/dawn simulators, and their effects on the human anatomy. Maan Van de Werken et al. [**7**] study artificial dawn and its effects on sleeping subjects, when exposed to light 30 minutes prior to waking. In particular, they studied the effects of exposure to light prior to waking on sleep inertia (grogginess), skin temperature, and cortisol levels. What Van de Werken found is that exposing heavy sleepers to artificial dawn prior to waking can reduce grogginess and increase activity. In addition, they found that skin temperatures showed a significant decline when waking in the artificial dawn condition versus the control condition. In another dawn simulation study, Thorn et al. [**11**] completed an experiment that focused on the change in cortisol levels when sleepers were exposed to light 30 minutes prior to alarm time. The experiment was conducted over the span of 4 consecutive weekdays, two days for a normal alarm and two days for their dawn-simulating *Natural Alarm Clock*. The participants were asked to take saliva samples upon awakening, which were collected each day in order to test their cortisol levels. Ultimately, Thorn et al. found that exposure to light prior to wakening led to a heightened cortisol response, when evaluated against cortisol responses to a traditional alarm clock.

3 SmartAlarm: iOS Development

I chose to develop *SmartAlarm* and the simpler *BasicAlarm* applications using *Android*. At the time, Android seemed to be the more appropriate Operating System because it enabled me to utilize my Java experience in development. However, HCI research pertaining to mobile application user studies emphasized that a functional prototype of the software is crucial for a valid evaluation **[1]**. Because most students that I know at Union College use Apple devices, I also learned *Swift*, Apple's programming language, in order to convert my applications from Android to iOS. Design choices for the iOS development of *SmartAlarm* can be seen in Appendix Section A.

While developing in iOS, I discovered a method for combining the functionality of *BasicAlarm* and *SmartAlarm* into one single *SmartAlarm* application. By including a light switch in the main scene of the application, I enabled users to select with ease between a typical, uni-sensory alarm notification (sound only) and a multi-sensory (sound and light) alarm notification. This would also make my usability study easier to conduct, because it would allow me to distribute one single application to my participants, rather than two. This was a crucial step in the development of my usability study, as I later discovered that simulating multiple applications on a single device is not supported by iOS.

3.1 iOS Notification Limitations

During the development of *SmartAlarm* in iOS, I came across several restrictions that apply to Apple's Operating System that do not apply to Google's. Primarily, iOS disables developers' ability to override the device's sound output for notifications. This is especially restricting for alarm clock application developers, because it requires that users enable their iPhone's sound output when setting their alarm. This leads to several usability challenges. First, without enabling the sound output, the alarm notification would lack sound, and would not wake sleeping users. Second, by enabling sound output, notifications from other applications, i.e. *iMessage* notifications, group chat application notifications, news applications notifications, etc., might wake the user if said notifications are not muted. Another challenge presented by iOS notifications is time restrictions. Apple limits mobile developers to a maximum notification duration of 30 seconds. The only way to avoid this problem is to automatically enable a repeating notification, which plays the notification's sound for 30 seconds in one minute intervals. However, I decided that implementing a repeating notification was unnecessary, as a 30 second sound would probably be enough to wake sleeping users. As stated earlier, the user can select between one of five sounds for their alarm. Because of iOS's notification duration limitation, I selected each of the five sounds based on their loudness and duration (each 30 seconds).

4 Multi-sensory Notification Implementation

The built-in capabilities of mobile devices disable a mobile developer's ability to produce a visual notification strong or bright enough to wake a sleeping user. Therefore, it was necessary to utilize third party software in order to implement a multi-sensory alarm notification. In this section, I describe the Philips Hue LED Lightbulb technology used in *SmartAlarm*, introduce the concept of "daylight simulation", and discuss my reasoning behind certain implementation details.

4.1 Philips Hue

Philips Hue LED Lightbulbs were designed to give users full control of their home's lighting directly from their mobile devices. Philips provides an open-source API for developing with Hue in both Android and iOS. However, lightbulb access and manipulation requires complete communication between the *four components* of the hue system:

1.) **Controller**: The software engineers behind Hue provided future developers with a Philips Hue API, which enables mobile developers to interact with the lightbulbs from their own software. With this comes the first component of the system, the **application**, or controller.

2.) **Portal**: The Philips Hue API uses a *RESTful interface* over HTTP. The purpose of the RESTful interface is to give each lightbulb a unique URL. Using API methods, HTTP commands (i.e. PUSH or GET commands) are sent to the lightbulb URLs, updating them accordingly.

A Philips Hue Lightbulb package comes with a Philips Hue Bridge and 1-3 Philips Hue lightbulbs (white or colored). These make up the third and fourth component of the system:

3.) **Bridge**: The Philips bridge is the centerpiece of the Hue system, and connects to a wireless router using a supplied Ethernet cable. Before proceeding to lightbulb customization, the application must discover and connect to a bridge. In order to do so, the mobile device using the application must be connected to the same wireless network as the bridge. Once connected to a bridge, developers can use the Hue API to send commands through the bridge to the lightbulbs.

4.) **Hue Lightbulbs**: The output of the system. Lightbulbs are updated by the RESTful interface and the bridge. Updates are controlled from the mobile application.

The flowchart below further explains the connection between the components listed above.

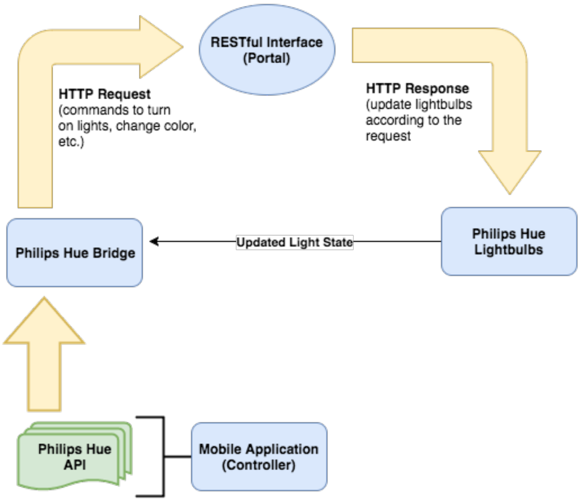


Figure 1: Flowchart of Philips Hue System

4.2 Daylight Simulation

An extensive amount of psychological and biological research has been conducted regarding daylight/dawn simulators, and their effects on the human anatomy and psyche. Maan Van de Werken et al. [7] study artificial dawn and its effects on sleeping subjects when exposed to light 30 minutes prior to waking. In particular, they studied the effects of daylight simulation on sleep inertia (grogginess), skin temperature, and cortisol levels. Van de Werken found that exposing heavy sleepers to artificial dawn prior to waking can reduce grogginess and increase activity. In addition, they found that skin temperatures declined significantly when waking in the artificial dawn condition versus the control condition. In another daylight simulation study, Thorn et al. [11] completed an experiment with specific focus on cortisol levels when sleepers were exposed to 30 minutes of light prior to alarm time. The experiment was conducted over the span of 4 consecutive weekdays, two days for a normal alarm and two days for their dawn-simulating *Natural Alarm Clock* [11]. The participants were asked to take saliva samples upon awakening, in order to test their morning-time cortisol levels. Ultimately, the researchers were able to obtain conclusive results from a four day experiment. Thorn et al. found that daylight simulation led to a heightened cortisol response, when evaluated against cortisol responses to a traditional alarm clock.

In the implementation of *SmartAlarm* I decided to present the multi-sensory notification as a daylight simulator. To do so, I linearly incremented lightbulb brightness and saturation, starting at 0% 30 minutes prior to alarm time, and ending at 100% at the selected alarm time. The figure below represents the daylight simulation implemented in *SmartAlarm*.

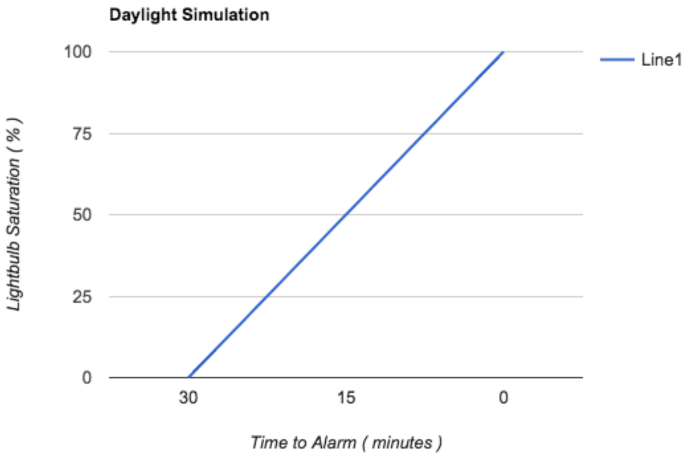


Figure 2: Daylight Simulation Implemented in SmartAlarm

5 Usability Study

To measure the effects that multisensory notifications have on user reactivity, I conducted a usability study on *SmartAlarm*. In this section, I present the rationale behind my choice of research methodology, the specific logistics of the usability study I performed, the information gathered, as well as the challenges and limitations presented by this study.

5.1 Research Methodology

Choice of research methodology is a crucial aspect of conducting an HCI usability study. With millions of apps leaving available in app stores on an annual basis, modern life seems to depend on the usability of mobile applications. However, user-based evaluation of mobile applications is still viewed as a relatively new branch of HCI research, creating a dearth of mobile usability study research among the HCI community. As a result, there is no concrete formula for conducting a user-based study for mobile applications, which forces flexibility in choice of research methodology. According to the limited mobile Human Computer Interaction research, choice of research methodology should be shaped to meet the needs of the study's research objectives [**1**].

HCI-based research on usability studies emphasizes two main methods for data collection, each consisting of specific advantages and disadvantages. *Laboratory studies* are studies in which participants are given instructions and are observed *directly*. While these studies are more conducive for collecting high-quality data regarding the usability of a mobile application, laboratory studies strip mobile applications of their real-life context [**1**]. In other words, laboratory studies have a tendency of limiting independent variables that might affect the application's usage. *Field studies*, on the other hand, are studies in which participants are given instructions and are observed *indirectly*. Data collection methods for field studies include online surveys, verbal recordings, and participant interviews. According to HCI research, field studies are a more efficient method for gathering data on user behavior and user satisfaction from a real-world context [**1**]. However, indirect observation creates an overwhelming reliance on honesty among participants. If participants are dishonest regarding the values of metrics recorded, study results can be incorrect or skewed without the researcher ever knowing.

In choosing the research methodology for this usability study, it was necessary to outline and prioritize the study's goals. The primary goal of this study was to measure the effects of the different notifications on user reactivity and behavior. A secondary goal was to measure user satisfaction and comfort when using my application. Because the analysis of user behavior and user satisfaction is more in-line with the goals of my study, I decided it was necessary for my main evaluation method to be a field study. Furthermore, because *SmartAlarm* requires wireless connectivity and because its users are asleep when the application is being used, it was necessary to provide participants with a real-life context of use. It is necessary to ensure that the application works as desired prior to conducting the field study. Therefore, I distributed the app to a group of people prior to conducting the field study.

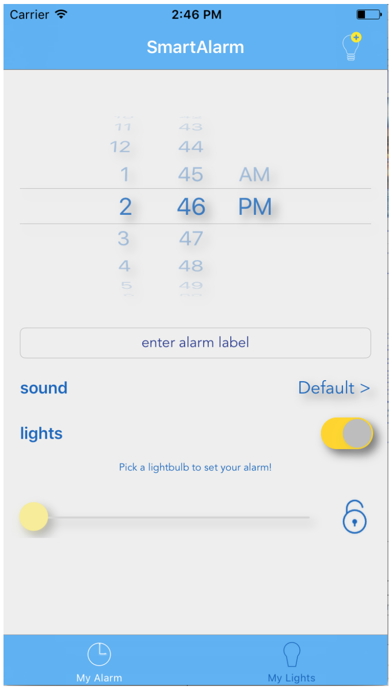
5.2 Preliminary Usability Study

I decided that it would be in my best interest to gather some feedback on SmartAlarm's usability, in order to further debug the application prior to conducting a field study. During the first week of the term, I distributed the app to a pool of three participants. Giving the participants very minimal instruction, I directly and individually observed them as they attempted to use the application, setting the alarm with both a uni-sensory and multi-sensory notification. This initial distribution of *SmartAlarm* allowed me to uncover multiple bugs that I was unable to observe with my ”developer bias”.

5.2.1 Accounting for Multiple Users

In testing and debugging *SmartAlarm* on my own, I failed to observe a significant bug in my implementation of the Philips Hue API. Originally, my code accessed all of the bulbs connected to a selected Philips Hue

Bridge when setting a multi-sensory alarm. Consequently, this changed all of the bulbs' light settings each time an alarm was set. Originally, this code worked flawlessly when I was the only person testing the alarm. In my study, however I planned on distributing multiple lightbulbs to different people using the same wireless network. In other words, when multiple users set alarms using the same Philips Hue Bridge, all of their lights would be activated upon the first alarm time. To prevent this issue, I conducted more research on the Philips Hue API, and discovered that single lights could be accessed using a unique identifier. I edited my implementation to include ”choose lightbulb” icon, displayed on the right-hand side of the top navigation bar the figure below. After the user selects their bridge and enables the light setting, this icon allows the user to select their specific lightbulb.



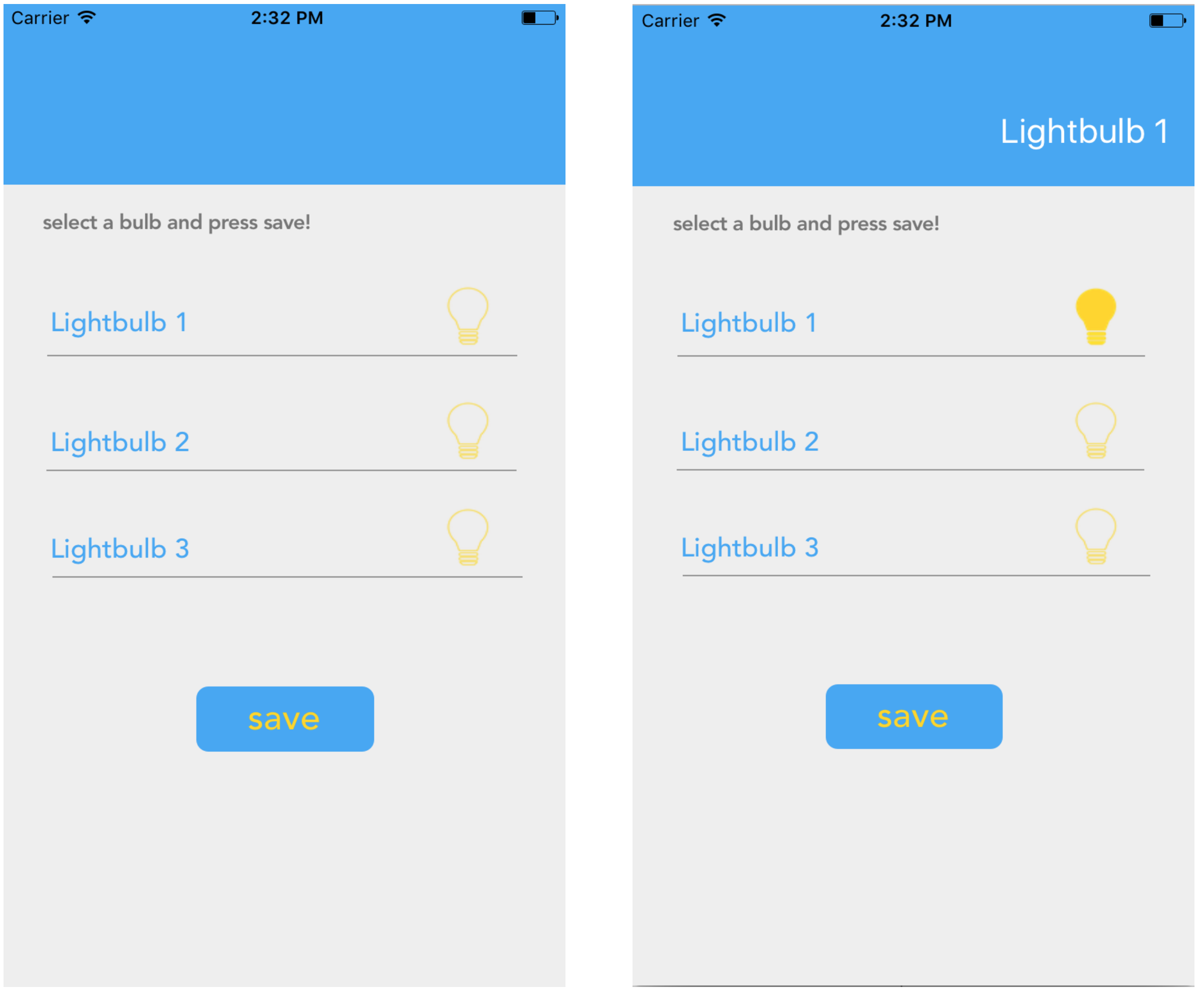
sound Default >

lights

Pick a lightbulb to set your alarm!

Figure 3: Lightbulb Selection Navigation Item

While the light setting is activated, *SmartAlarm* disables the user's ability to set their alarm until they select their specific lightbulb. The user is prompted to select a lightbulb with the text above the alarm slider. Clicking the navigation item presents the user with the following scene.



select a bulb and press save!

select a bulb and press save!

(a) Prior to Selecting a Lightbulb

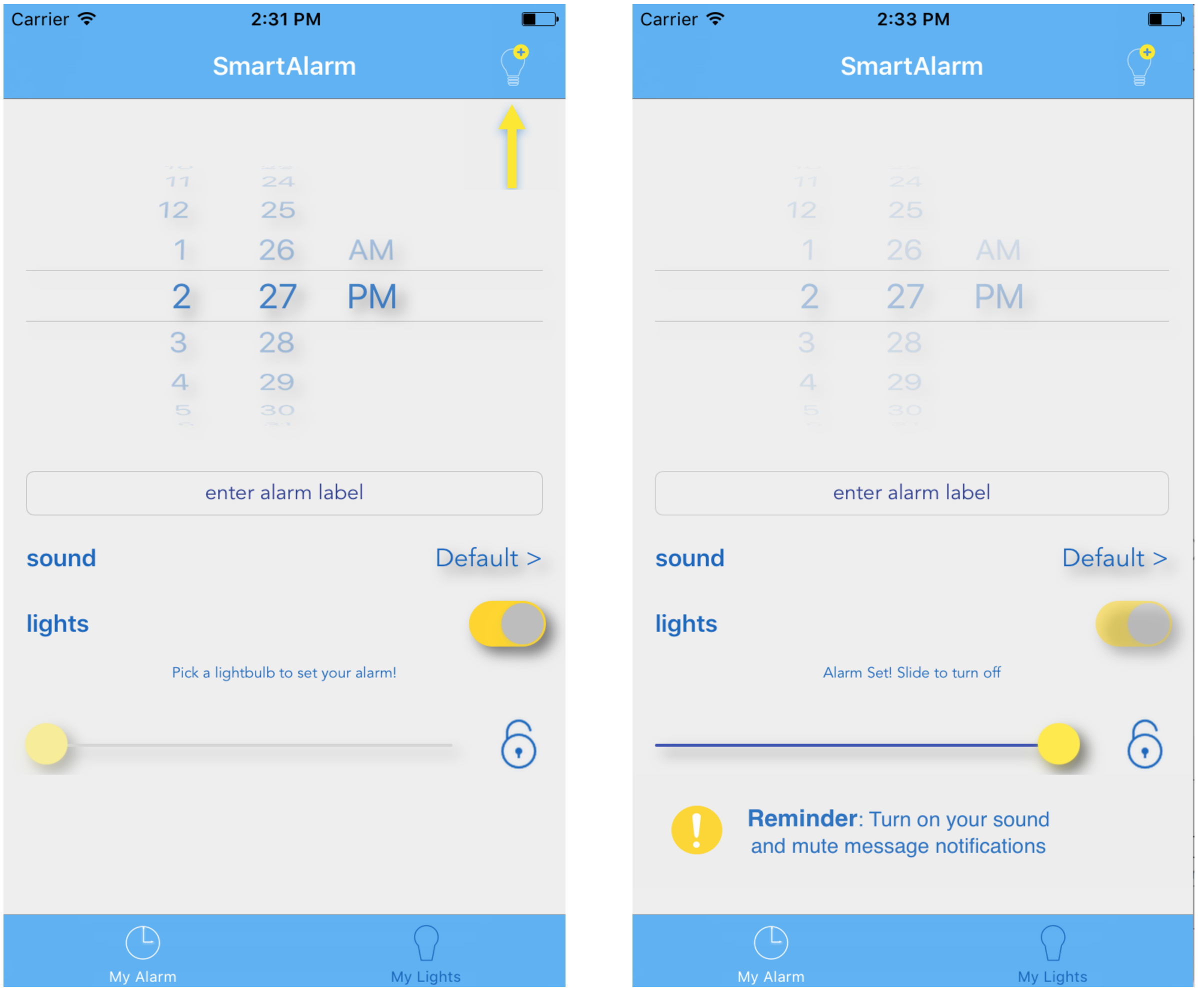
(b) After Selecting a Lightbulb

Figure 4: Lightbulb Selection Scene Before and After Selecting

Selecting one of the lightbulbs briefly flashes that lightbulb, notifying the user that they have chosen the correct bulb.

5.2.2 Instruction Suggestions

After fixing the bulb selection bug, I had my participants test my alarm for one night using the light setting. When running the app simulation on their phones, I verbally instructed them to activate their devices sound output. While one of the users was able to use the app without further instruction, the other two suggested that I add more visual instructions within the application, for selecting a lightbulb and for turning on their device's sound output. Upon request, I added the following prompts to further instruct the user.



Pick a lightbulb to set your alarm!

Alarm Set! Slide to turn off

Reminder: Turn on your sound and mute message notifications

(a) Light Selection Prompt (b) Sound Output Prompt

Figure 5: User Suggested Instructions

Figure 9a shows the light selection prompt that is displayed when the user enables the light setting. I decided that the arrow combined with the text above the slider should be enough to instruct the user to perform this action. Figure 9b displays the instruction that is presented to the user after setting their alarm, prompting them to enable their device's sound output and mute other notifications.

5.3 Experimental Design

To measure the effects that multi-sensory notifications have on user reactivity and behavior, I conducted a field study on a total of 34 Union College Seniors. The study was conducted over a 6 week period, in which each of my participants spent four days waking with *SmartAlarm*. On Sunday and Monday nights, participants were instructed to set their alarms using multi-sensory notifications. On Tuesday and Wednesday nights, participants were instructed to set their alarms using typical uni-sensory alarm notifications, disabling the light setting. I decided two nights each would be appropriate for collecting data, as related studies were able to retrieve conclusive results using this time-frame [**11**]. In addition, because time was limited due to the structure of this senior project, and resources were limited due to the amount of hardware I had, I felt it was necessary to diversify my pool of participants with shorter time-frames for each participant.

Prior to conducting the study, I had my participants fill out a preliminary sleep habit-based survey, which measured their Sleep Quality Index (SQI) as described in Section 2. The SQI takes 8 sleep-based factors into account, and ranks each factor on a scale of 0 to 2. Once rankings for all factors are reported by the participants, the rankings are summed to calculate each participant's SQI. The factors taken into account by the SQI are displayed in Table 1. By recording participants' Sleep Quality Indexes, I had hoped to differentiate between good sleepers and poor sleepers, and draw some conclusions on the effects of daylight simulation on different types of sleepers.

As stated earlier, participants were instructed to spend two days waking with daylight simulation, followed by two days waking with a typical alarm clock notification. During the mornings following their use of *SmartAlarm*, participants were instructed to fill out an Internet-based survey, which recorded the following metrics:

1. Original alarm time,
2. Reaction time (time the user actually got out of bed),
3. Number of snooze presses,
4. Level of comfort waking with SmartAlarm (from 1 to 5, 1 being not comfortable),
5. Level of morning-time grogginess (from 1 to 5, 1 being not groggy)

After users reported this information, user reactivity was derived by taking the difference between reaction time and original alarm time in minutes. I decided that comparing the data received for uni-sensory notifications against data received for multi-sensory notifications would give me enough information to provide some answers to the questions addressed in Section 1. Prior to conducting the study, I had a notion that the last performance measure, participants' level of grogginess, would be a relative and potentially invalid measure, because college students might be tired in the morning due to a variety of reasons. However, I felt it was necessary to record this information, to see if daylight simulation had any clear effects on tiredness.

5.4 Challenges and Limitations

There were a plethora of challenges in conducting this specific usability study. The Philips Hue technology used to provide the user with a multi-sensory alarm experience presented a number of limitations. Primarily, Philips Hue Lightbulb connection requires an Ethernet port that connects to the same wireless network as the user's mobile device. Due to issues presented by Union's on-campus wireless network, my participants were limited to Union College Seniors living in off-campus houses. While this pool of participants suggests a potential of biased results, I did not suspect that the limitation would hinder my ability to collect a sufficient amount of data. Another limitation was due to the amount of hardware I had to conduct the study (2 Philips Hue Bridges, 6 Philips Hue Lightbulbs). As a result, my study was limited to 6 field tests per day.

Using an indirect observation method presented another significant challenge. Because the study was conducted in the form of a field test, participants were trusted to complete the given instructions, such as actually setting the alarm, noting their wake up time, filling out surveys, etc. At first, I used email as my method for sending out the morning-time surveys. However, I found myself having to remind many of my participants to fill out the surveys. To fix this issue, I started sending the surveys to participants via iMessage, which seemed to be a more convenient data collection method for my participants.

Another challenge presented by this study resided in limiting independent variables. Some of the independent variables presented by this study included room arrangement, window placement/the amount of outside light in the bedroom, and the distance between participants' head while sleeping and the positioning of the lightbulb. In order to limit these independent variables, I asked permission to set up participants' bedrooms prior to conducting the study. In some cases, I felt it was necessary to cover participant windows with black-out shades to eliminate the outside light factor. However, many of my participants already had curtains on their windows to eliminate this factor. To be sure that each participant receives the same notification, I made sure that the light was positioned at least 3 feet away from the edge of the participant's bed. Furthermore, some of my participants did not have a lamp, and I had to provide one of my own for their use.

Perhaps the most significant limitation I faced in conducting this study was the time constraint presented by the structure of this project. Because the study was limited to only 7 or 8 weeks, I was forced to make some tough decisions regarding participation, and analyze the tradeoffs between my options. While conducting a longer study on fewer participants might have produced more accurate results, I felt that asking college seniors to spend more than 4 days participating would be too much to ask with no incentive. Also, having fewer participants might have provided biased results. Therefore, I decided to diversify my pool of participants by limiting participation to a weekly basis.

6 Data Analysis

I completed the study with a combined total of 134 instances of user reactivity data – 67 instances of uni-sensory data and 67 instances of multi-sensory data. Missing values were expected, because I suspected there would be certain cases in which participants forgot to set their alarm, or forgot to submit their data. However, only four instances of reactivity data were not submitted. Fortunately, two instances of uni-sensory data and two instances of multi-sensory data were missing, and the amount of data for uni-sensory notifications and multi-sensory notifications remained equivalent after missing values were accounted for.

In this section, I provide some high-level takeaways from my data analysis, followed by a more in-depth analysis of the data received for each of the performance metrics recorded.

6.1 High-Level Takeaways

The usability study provided me with some conclusive results regarding each of the metrics recorded in my morning-time surveys. Primarily, my findings showed that daylight simulation had the most significant effect on reported grogginess when the snooze button was not utilized by participants. When the snooze functionality of *SmartAlarm* was not used, users were more likely to report low grogginess levels when presented with daylight simulation prior to waking. On the other hand, users were much more likely to report higher grogginess levels when presented with typical alarm notifications. Without isolating the instances of grogginess data based on snooze presses, I did not trust grogginess as an effective measure to be significantly affected by the different notifications. Pertaining to reactivity, I found that participants were somewhat more likely to react quickly to a multi-sensory alarm clock notification. On the contrary, participants were more likely to report a slower reaction when presented with a uni-sensory alarm clock notification. With respect to the "snooze presses" metric, participants appeared to be **more reluctant** to press the snooze button with multi-sensory notifications. Conversely, users were significantly more likely to press snooze 3 or more times with sound as their only method for waking. Regarding user comfort, I discovered when using a typical alarm clock notification, participants were much more likely to report the highest level of comfort waking with *SmartAlarm*.

6.2 Daylight Simulation and Reported Grogginess

Multi-sensory alarm clock notifications, or daylight simulation notifications, proved to have the greatest effect on reported morning-time grogginess when the snooze button was not pressed. As described in Section 4.2, prior studies conducted on the effects of daylight simulation on the human anatomy and psyche began the daylight simulation process 30 minutes prior to alarm time [**7**] [**11**]. In these studies, participants were not given the choice to get more sleep after the original alarm time. After the snooze button is pressed in *SmartAlarm*, the light notification is presented at the same time the sound notification is presented, not before. Therefore, in order to observe the actual effects of daylight simulation on grogginess, I isolated the instances of data in which the snooze feature was not utilized by participants. Based on the findings presented by previous daylight simulation studies, my results make sense in that reported grogginess levels for multi-sensory notifications tended to be lower when snooze was not pressed. My findings suggest that daylight simulation loses its effects on grogginess when users choose to get more sleep after the initial alarm.

A total of 50 instances of data were collected when the snooze feature was not used – 31 multi-sensory instances and 19 uni-sensory instances. To make the following discussion simpler, I define low grogginess levels as reported grogginess levels of 1 and 2, and high grogginess levels as reported grogginess levels of 3, 4, and 5. For the multi-sensory instances, roughly 68% of participants reported low grogginess levels, and about 32% reported higher grogginess levels. For the uni-sensory instances, about 42% of users reported low grogginess levels, while roughly 58% of users reported high grogginess levels. These results are visually represented in the table and bar graph below.

Notification Type	**Low Grogginess**	**High Grogginess**
Multi-sensory	$\approx 68\%$	$\approx 32\%$
Uni-sensory	$\approx 42\%$	$\approx 58\%$

Table 2: Notification Type vs. Reported Grogginess with No Snooze Presses

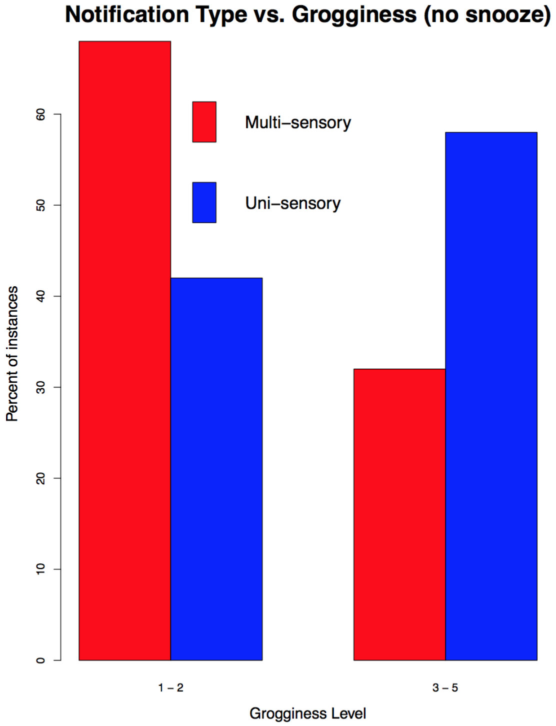


Figure 6: Bar Graph of Notification Type vs. Reported Grogginess with No Snooze Presses

Based on these results, I was able to conclude that use of the snooze button limits the effects of daylight simulation observed prior studies. When the snooze button is not used, participants were roughly 62% more likely to report low grogginess levels with multi-sensory notifications. Moreover, without using the snooze button, participants were about 81% more likely to report high grogginess levels with sound as their only method for waking.

6.3 User Reactivity Analysis

To provide a clear analysis of user reactivity data against notification type, I grouped instances of user reactivity data into three categories. I define ”quick reactions” as reactions under 10 minutes, ”average reactions” as reactions between 10 and 29 minutes, and ”slow reactions” as reactions greater than or equal to 30 minutes. While there were no significant results with respects to average reactions, I was able to come to some conclusions regarding quick reactions and slow reactions. For the multi-sensory instances, about 35% of participants reported quick reactions, while about 9% of participants reported slow reactions. For the uni-sensory instances, about 29% of users reported quick reactions, while about 18% of users reported slow reactions. These findings are visually represented in the table and figure below.

Notification Type	**Quick Reactions**	**Average Reactions**	**Slow Reactions**
Multi-sensory	$\approx 35\%$	$\approx 56\%$	$\approx 9\%$
Uni-sensory	$\approx 29\%$	$\approx 53\%$	$\approx 18\%$

Table 3: Notification Type vs. User Reactivity

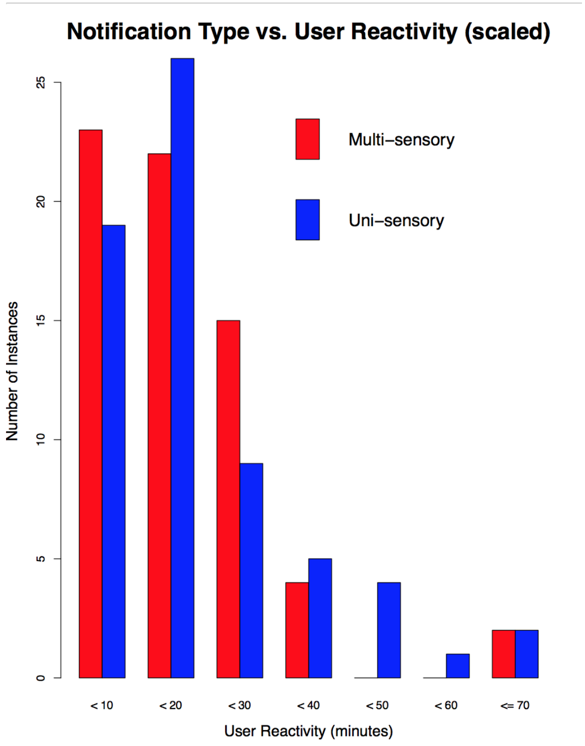


Figure 7: Bar Graph of Notification Type vs. User Reactivity

My findings enabled me to draw a couple conclusions regarding user reactivity and notification type. Primarily, I found that users were roughly 21% more likely to react to their alarm in under 10 minutes when presented with multi-sensory notifications. Furthermore, my findings show that participants were roughly 50% more likely to have slow reactions when sound was their only method of waking.

6.4 Snooze Usage Analysis

Upon analyzing the snooze metric, participants appeared to be more reluctant to use the snooze feature when the light setting of *SmartAlarm* was activated. These results confirm my findings pertaining to daylight simulation and morning-time grogginess – if users tend to be more active when presented with multi-sensory alarm notifications, they will probably be less likely to desire more sleep after their initial alarm. While the data for one or two snooze presses was inconclusive, participants were significantly more likely to press the snooze button three or more times with sound as their only method for waking. These results tell me that participants tended to be lazier or more negligent of their initial alarms when presented with uni-sensory alarm clock notifications. The results for the snooze metric are visually represented below.

Notification Type	**0**	**1**	**2**	**3**	**4**	**5 or more**
Multi-sensory	$\approx 41\%$	$\approx 27\%$	$\approx 24\%$	$\approx 4.5\%$	0%	$\approx 3\%$
Uni-sensory	$\approx 29\%$	$\approx 39\%$	$\approx 14\%$	$\approx 9\%$	$\approx 4.5\%$	$\approx 4.5\%$

Table 4: Notification Type vs. Number of Snooze Presses

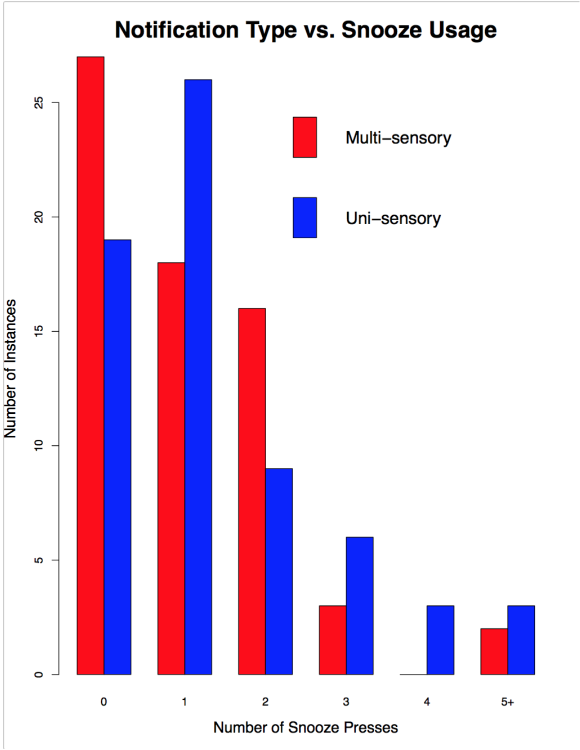


Figure 8: Bar Graph of Notification Type vs. Number of Snooze Presses

Based on the data above, I was able to conclude that participants were roughly 41.4% more reluctant to press the snooze button with a multi-sensory alarm clock experience. Furthermore, users were roughly 140% more likely to press the snooze button three or more times with a uni-sensory alarm clock experience. The results for one and two snooze presses are inconclusive, as participants were about 44% more likely to press snooze once with sound as their only method for waking, and $\approx 71.4\%$ more likely to press snooze twice when waking with both sound and light as their waking method.

6.5 User Comfort Analysis

Most of the data reported on user comfort when waking with *SmartAlarm* was relatively inconclusive, however, I found that while only 9% of participants reported the highest level of comfort for the multi-sensory instances, $\approx 23\%$ of participants reported the highest level of comfort when waking with uni-sensory notifications. The data for reported user comfort is displayed in the table and bar graph below.

Notification Type	**1**	**2**	**3**	**4**	**5**
Multi-sensory	$\approx 3\%$	$\approx 6\%$	$\approx 39\%$	$\approx 42\%$	$\approx 9\%$
Uni-sensory	$\approx 3\%$	$\approx 3\%$	$\approx 29\%$	$\approx 42\%$	$\approx 23\%$

Table 5: Notification Type vs. User Comfort

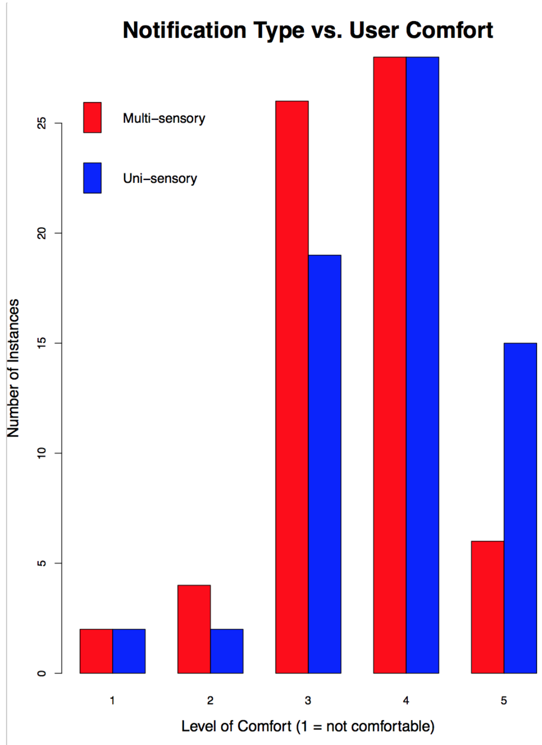


Figure 9: Bar Graph of Notification Type vs. User Comfort

Participants were $\approx 155\%$ more likely to have the highest level of confidence in *SmartAlarm* after waking with uni-sensory notifications. On the other hand, participants were about 37.1% more likely to report low comfort levels (1, 2, and 3) with multi-sensory notifications. The higher level of user confidence in a typical alarm clock notification may be a result of people being more comfortable with the things they are used to. Another factor that may have influenced these results is that users might be less comfortable with using foreign technology that was implemented by an inexperienced mobile developer.

6.6 Reported Grogginess Analysis

Prior to conducting the study, I suspected that grogginess would be a relative measure because my participants consisted of a pool of college students. Knowing this, it was necessary to take the data received for this metric with a grain of salt, as college seniors could feel tired in the morning for a variety of reasons. However, upon analyzing this data I noticed a trend – users tended to report higher grogginess levels when presented with uni-sensory notifications as opposed to multi-sensory notifications.

Notification Type	**1**	**2**	**3**	**4**	**5**
Multi-sensory	$\approx 11\%$	$\approx 47\%$	$\approx 26\%$	$\approx 12\%$	$\approx 4.5\%$
Uni-sensory	$\approx 16\%$	$\approx 23\%$	$\approx 38\%$	$\approx 18\%$	$\approx 4.5\%$

Table 6: Notification Type vs. Reported Grogginess

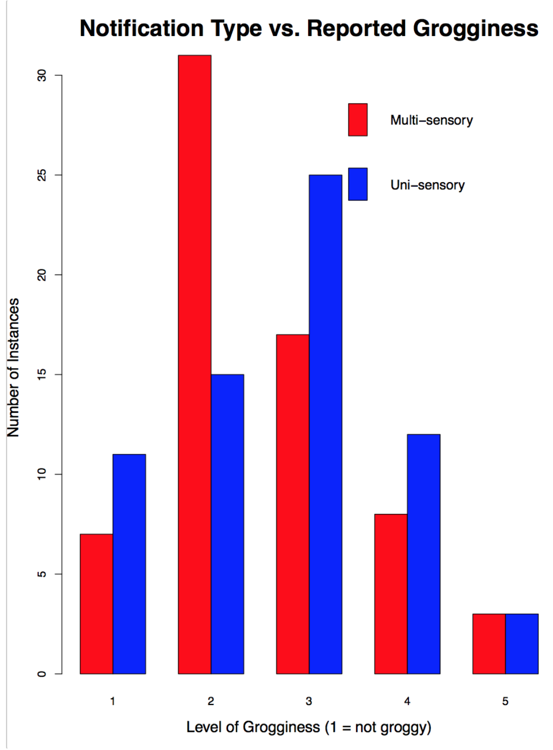


Figure 10: Bar Graph of Notification Type vs. Reported Grogginess

Based on these results, users appear to be $\approx 42.4\%$ more likely to report higher grogginess levels (3, 4, and 5) when presented with sound as their only method for waking. However, while these results may seem conclusive, they could also be coincidental due to the relativity of the measure.

6.7 Insignificant Results

Upon cross-analyzing my results from the Sleep Quality Index Survey and user reactivity, I did not notice any significant trends. In fact, the reactivity data received for good sleepers against bad sleepers appeared to be the opposite of my expectations. While I had hoped that daylight simulation would be the more effective method in helping ”bad sleepers” wake quickly, the bad sleepers were actually more likely to react quickly with uni-sensory alarm notifications.

7 Conclusion

Prior to the mobile application movement of the 21st century, research regarding desktop-based usability studies made up a large portion of HCI research. However, the mobile movement has created a demand for change and flexibility in the HCI research community. Furthermore, as notifications become an integral aspect of all mobile applications, it is necessary for the HCI research community to focus on mobile notification effectiveness and usability. In conducting this study, I had hoped to answer several questions regarding multi-sensory alarm clock notifications and their effects on a user's alarm clock experience. My primary objective was to determine whether or not simulating light exposure prior to waking could heighten user reactivity. Based on my findings from this relatively limited study, I was able to conclude that users were more likely to react quickly when presented with a multi-sensory alarm clock experience. Another question I aimed to address was whether or not daylight simulation could reduce morning-time grogginess. While the grogginess measure was relative, I was able to analyze the actual effects of daylight simulation by isolating the instances of data in which the snooze feature was not used. After doing so, I determined that users were less likely to report high morning-time grogginess levels when presented with daylight simulation prior to waking. These results align with the results provided by previous biological and psychological studies conducted on daylight simulation and its effects on the human anatomy and psyche **[7]** **[11]**. Furthermore, I had hoped to answer whether users are more or less comfortable waking with a multi-sensory notification. My findings showed that users are, in fact, more comfortable waking with typical alarm clock notifications. This suggests that users are more comfortable with what they are used to, or that users are less comfortable with using an atypical technology that was developed by a novice mobile developer. Finally, while I had hoped to draw some conclusions between different types of sleepers and reactivity, my results on the matter were inconclusive. All in all, while the study conducted was limited by several pre-determined factors, I was able to provide some answers to the questions I had originally hoped to address.

8 Future Work

The observed effectiveness of multi-sensory alarm notifications, as well as the implementation of the multi-sensory notification technology in *SmartAlarm*, provides a decent amount of potential for future research. Given more time to conduct the study, I would extend this study by a month or two, in order to increase the amount of data for each participant. The hope is that I would have enough information to compare results received for individual participants with this data, to see how switching between a uni-sensory and multi-sensory can affect an individuals sleep habits.

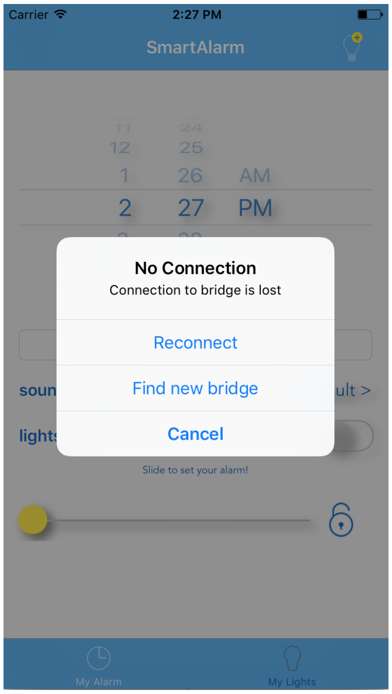
The multi-sensory notifications implemented in *SmartAlarm* can be used for more than just alarm clock notifications. For example, these notifications can be used in future research for notifying the deaf of important events, or remind them to complete daily tasks. While gathering a pool of deaf participants to conduct this study might be a difficult task, this technology could potentially be used to help them live better lives.

Appendices

A SmartAlarm: iOS Design Choices

Figure 11: SmartAlarm: iOS App Icon

Upon entering *SmartAlarm* for the first time, the user is prompted to select their *Philips Hue Bridge*, in order to access and manipulate their Philips Hue lightbulb.



No Connection
Connection to bridge is lost

Slide to set your alarm!

Figure 12: Philips Hue Bridge Connection Prompt

Once a bridge is selected and confirmed, the bridge is stored in the application's "user defaults", so that the user does not need to reconnect every time the application is closed and re-opened. The light switch is disabled until the user selects their bridge.

A.0.1 My Alarm Tab

The first and most important functionality of *SmartAlarm* is the alarm itself. In the *My Alarm* tab, the user can customize and set their alarm. The design for this scene is quite minimalistic, allowing users to pick a time, set a label to attach to their alarm notification, pick one of five sounds, and enable or disable the application's light setting. The scene for this tab is displayed in the figure below.

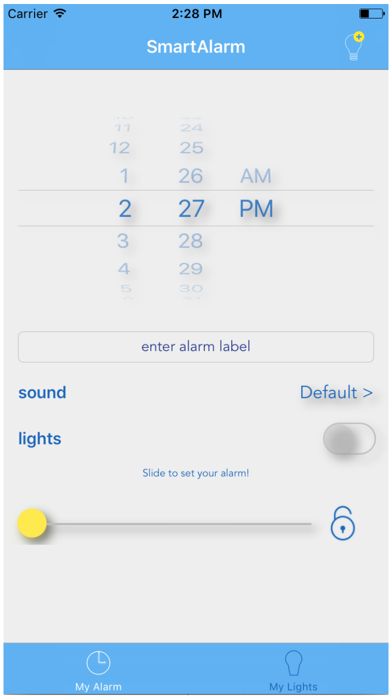


Figure 13: My Alarm Scene

To set the alarm, I decided to implement the operating system's built-in slider, to ensure that the user would not accidentally set the alarm to undesirable preferences. To instruct the user on setting the alarm, I decided to implement a dynamic prompt above the slider. The instructions change as the user interacts with the slider, as shown in the figure below.

(a) Alarm Slider and Turn On Prompt

(b) Alarm Slider and Turn Off Prompt

Figure 14: Alarm Slider and Prompt Before and After Setting

The lock icon on the right-hand side of the slider allows the user to temporarily disable the alarm slider, so that they don't unintentionally turn the alarm off. The icon is dynamically changed, based on the user's actions. This functionality is displayed in the figure below.

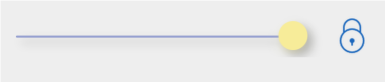


Figure 15: Lock Button Functionality

A.0.2 My Lights Tab

In development, I decided to add more functionality to *SmartAlarm*, enabling the user to control their Philips Hue Lightbulb. The *My Lights* tab allows the users to turn on and off their lightbulb, and to select between five different pre-designated light settings. The figure below displays the scene for this tab.

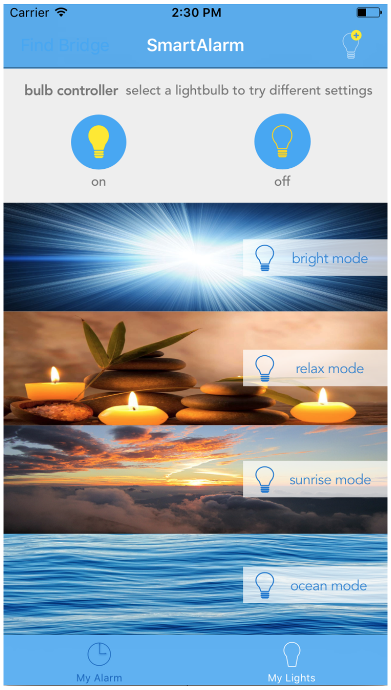


Figure 16: My Lights Scene

I felt it was necessary to add this functionality within my application, so that users wouldn't need to outsource to another Philips Hue Developer to control their lightbulbs.

References

[1] Alshehri, Fayez and Freeman, Mark. *Methods of usability evaluations of mobile devices*. 23rd Australasian Conference on Information Systems, Pages 1-10, 2012.

[2] Booker, John E. and Chewar, C. M. and McCrickard, Scott D. *Usability Testing of Notification Interfaces: Are We Focused on the Best Metrics?*. `http://people.cs.vt.edu/mccricks/papers/bcm04.pdf` Proceedings of the 42nd Annual Southeast Regional Conference (ACM-SE 42), pages 128-133, 2004.

[3] Buboltz, Walter C. and Brown, Franklin and Soper, Barlow. *Sleep Habits and Patterns of College Students: A Preliminary Study*. `https://www.researchgate.net/publication/1593058` Journal of American College Health, Volume 50 Number 3, pages 131-135, 2001.

[4] Choe, Eun Kyoung and Fonville, Amanda and Kientz, Julia A. and Sakaguchi, Dawn and Halko, Sajanee and Watson, Nathaniel F. *Opportunities for Computing to Support Healthy Sleep Behavior*. `http://designsquare.ist.psu.edu/download/CHI-10-Choe.pdf` Association for Computing Machinery, pages 3661-3666, 2010.

[5] Alshehri, Fayez and Freeman, Mark. *Methods of usability evaluations of mobile devices*. `https://dro.deakin.edu.au/eserv/DU:30049129/alshehri-methodsofusability-2012.pdf` 23rd Australasian Conference on Information Systems, pages 1-10, 2012.

[6] LeeTiernan, Scott and Cutrell, Edward and Czerwinski, Mary and Hoffman, Hunter. *Effective Notification Systems Depend on User Trust*. `http://research.microsoft.com/en-us/um/people/marycz/trust01.pdf` Microsoft Research, 2001.

[7] Van de Werken, Maan and Gimenez, Marina and De Vries, Bonnie and Beersma, Domien and Van Someren, Eus J. W. and Gordijn, Marijke C. M. *Effects of artificial dawn on sleep inertia, skin temperature, and the awakening cortisol response*. `http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2869.2010.00828.x/epdf` Journal of Sleep Research, Volume 19, pages 425-435, 2010.

[8] McCrickard, Scott D. and Czerwinski, Mary and Bartram, Lyn. *Introduction: design and evaluation of notification user interfaces*. `http://people.cs.vt.edu/mccricks/papers/mcb03.pdf` International Journal of Human-Computer Studies, Volume 58, pages 509-514, 2003.

[9] McCrickard, Scott D. and Catrambone, Richard and Chewar, C. M. and Stasko, John T. *Establishing tradeoffs that leverage attention for utility: empirically evaluating information display in notification systems*. [*International Journal of Human-Computer Studies, Volume 58, pages 547-582*]. `http://www.cc.gatech.edu/ stasko/papers/ijhcs03.pdf` Elsevier Science Ltd, 2003.

[10] Statista. *Number of mobile app downloads worldwide from 2009 to 2017 (in millions)*. `https://www.statista.com/statistics/266488/forecast-of-mobile-app-downloads/` Statista – The Statistics Portal, 2016.

[11] Thorn, L. and Hucklebridge, F. and Esgate, A. and Evans, P. and Clow, A. *The effect of dawn simulation on the cortisol response to awakening in healthy participants*. `http://www.lampe-de-luminotherapie.com/wp-content/uploads/effets-luminotherapie.pdf` Elsevier: Psychoneuroendocrinology, Volume 29, pages 925-930, 2003.

[12] Zhang, D. and Adipat, B. *Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications*. `https://pdfs.semanticscholar.org/ca91/2a18f14df60db103abaf156d4bba2c50e26c.pdf` International Journal of Human-Computer Interaction, Volume 18 Number 3, pages 293-308, 2005.