Designing Personal Systems for Mindful Decision Making

by

Niaja Nichole Farve

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Signature redacted

Author ....
Department of Electrical Engineering and Computer Science
May 18, 2016

Signature redacted

Certified by ....
Pattie Maes
Professor of Media, Arts and Sciences
Thesis Supervisor

Signature redacted

Accepted by ....
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Theses

Designing Personal Systems for Mindful Decision Making

by

Niaja Nichole Farve

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Computer Science and Engineering

Abstract

Today's personal technologies are generally seen as reducing mindfulness. Users are so absorbed in their devices that they behave in more distracted ways, are less engaged in face-to-face social interactions and increase their sedentary behaviors. This often results in behaviors and habits that are misaligned with the user's goals. Current attempts to use technology to improve well-being, such as fitness trackers, do not take advantage of some of the benefits that mobile, personal technologies have to offer. Specifically, increasingly mobile personal technologies have the opportunity to intervene in the moment when a person is making a decision with personalized, "just-in-time" nudges that may result in a more mindful decision. This thesis explores how to design personalized, wearable technologies that can support more mindful behavior. It investigates the various challenges that exists when designing such systems and provides design considerations for future systems. Human behavior researchers have argued that although a user may have the motivation and the ability to change behavior, a trigger is required to make a new behavior happen. This thesis specifically focuses on considerations that should be made when designing triggers for persuasive, wearable systems. These include ensuring the user's attention, utilizing contextual cues to determine timing of triggers and using personalized messages in a trigger. The thesis presents several pilots studies in using personal, wearable technologies to offer "just-in-time" triggers for behavior. The design and implementation of these systems is detailed and preliminary data regarding their effectiveness is discussed. These systems explore what challenges emerge when applying traditional behavior change theories on personalized, wearable systems.

Thesis Supervisor: Pattie Maes
Title: Professor of Media, Arts and Sciences

Acknowledgments

"The human spirit must prevail over technology." -Albert Einstein

"To be good and to do good is all we have to do." -John Adams

MIT is not a place where you can succeed alone. Without friends and family that have supported and motivated me, completing this thesis and succeeding at MIT would not have been possible. This includes friends before MIT and friends I have met at MIT.

Several groups on campus have served as extended families. This includes BGSA, ACME, Chocolate City, ODGE and many more. Without which, MIT would not feel like home.

I am especially grateful for my mother, role model, and best friend. You have pushed me to try even when success seemed impossible. You have always been there to listen to my problems and ensure me that I can overcome an obstacle. I hope to one day become half the mother you have been to me.

MIT would not have been as enjoyable without the man in my life that has been unbelievably understanding and encouraging. Thank you Omar for supporting my late nights and trying to understand my frustrations as I navigate the maze that is MIT.

Without my adviser Pattie Maes, this thesis would of course not been possible. Thank you for taking a chance and accepting me into your group. Having the ability to follow my own path when it came to my research made the experience immensely more enjoyable.

Finally, thank you to the Robin Wood Johnson Foundation and the Well-being SIG at the MIT Media Lab for supporting my research. Their financial support allowed me to run user studies and purchase materials.

This doctoral thesis has been examined by a committee of the Department of Electrical Engineering and Computer Science, Media Arts and Sciences and Sloan School of Management as follows:

Professor Pattie Maes ....

Chairman, Thesis Committee

Program Head, Media Arts and Sciences

Professor Renee Gosline ....

Member, Thesis Committee

Assistant Professor, Sloan School of Management

Professor John Guttag ....

Member, Thesis Committee

Dugald C. Jackson Professor of Electrical Engineering and Computer Science

Professor Peter Szolovits ....

Member, Thesis Committee

Professor of Electrical Engineering and Computer Science

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The growth in computational power has resulted in pervasive technology that produced countless advances. The average person reaps these benefits in the form of increased productivity and convenience. These benefits have not come without tradeoffs, however. The increased dependence on digital technology and mobile devices has lead to a rise in detrimental behaviors, including an ever-growing increase in sedentary lifestyles and decreases in face-to-face social interactions. Dropping these newly acquired behaviors has proved to be increasingly difficult, as their regularity has resulted in habit formation. Habits, behaviors that are regular in tendency and hard to give up are fueled by the habit loop as seen in Figure 1-1 [2] where known or unknown triggers prompt us to perform an action which has an immediate reward, but may be undesired when trying to adopt a healthy lifestyle. Changing or replacing these habits requires conscious attempts to avoid their triggers or replace their meaning.

The digital age has introduced an overwhelming number of devices that are competing for our attention and add to our cognitive load. This has resulted in a constant battle to tackle the cognitive demands of day-to-day life, technology and habit change. Typically our cognitive load is depleted before we can tackle habit change. Technology has therefore received a backlash and been identified as one of the main culprit for the current decline in health. Regardless, this has not stopped the rise of technology, which has not only become ubiquitous but also more interwoven into our lives and bodies. Therefore, it is not surprising that the quantified self movement has formed with people attempting to use technology to quantify their behaviors and learn from them [3]. As technology becomes more capable of monitoring behaviors, a fundamental shift can occur in which decision-making can be supported by technology. This shift could not only help to relieve the cognitive load, but could also assists in behavior change and new habit formation.

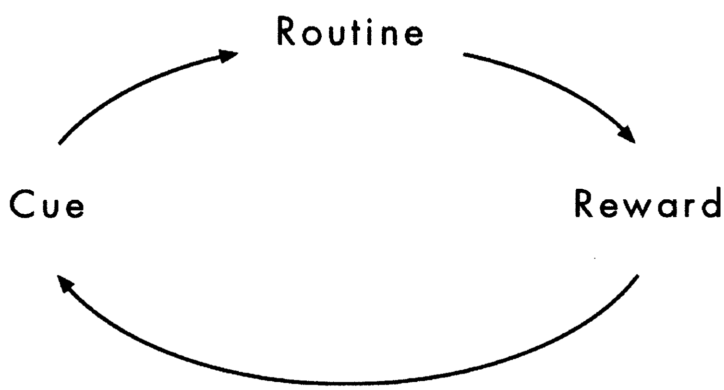


Figure 1-1: Habit loop by Charles Duhigg

1.1 Behavior Change

One of the most used frameworks for behavior change is BJ Fogg's model [1], which provides a simple formula with key variables for thinking about behavior change. The three variables, motivation, ability and trigger are all needed when producing a behavior. However, simply having each variable present does not guarantee a behavior to occur. As described in 4-3 the degree of motivation and ability are the key factors for a behavior to happen, where it is most desired to have someone be highly motivated and the behavior be easy to perform. After motivation and ability are optimized a trigger is needed as the final factor to producing a behavior.

When using this model to design persuasive systems, 3 strategies are taken to ensure behavior change. First, attempts can be made to increase a user's motivation.

This could mean providing the user new information that aligns with their personal motivations and goals or providing an extrinsic motivation such as the option to earn a prize. Second, the desired behavior should be made as easy as possible to perform. This could mean providing the user with the necessary resources or support. A trigger is provided to remind the user to execute the desire behavior.

If we imagine a user who recently committed to eating healthier, balanced meals, we can envision a system that helps with their goal by introducing a new behavior. This user visits the same restaurant every day to have a bacon cheeseburger for lunch. The restaurant serves as a trigger for the bacon cheeseburger. Personal technologies could help the user change their habit by presenting information in real time. The persuasive system could increase the motivation of the user to eat healthy by reminding them of their goal and perhaps emphasizing the undesired health effects of their current choice. It could also make it easy for the user to eat something else maybe by pointing out other healthy options in the restaurant. The user is now re-motivated to make healthy food choices, has an easy way to move toward that goal or is rewarded with the satisfaction of making a healthier choice.

As we can see from this example, personal systems could be effective behavior change tools. They can influence the user just before or at the time that a user makes a behavior choice. Additionally, they have the opportunity of being optimized for the user's wants and personal traits. With recent advances in technology (for example machine learning), they also have the ability to learn from our actions and adapt to new situations or even changing priorities. Finally, their ubiquitous nature allows them to be with the user at all times. However, defining effective personal systems to induce behavior change is not as simple as using the context to increase motivation and making the behavior easy. There are several other factors that need to be considered when designing such personal systems.

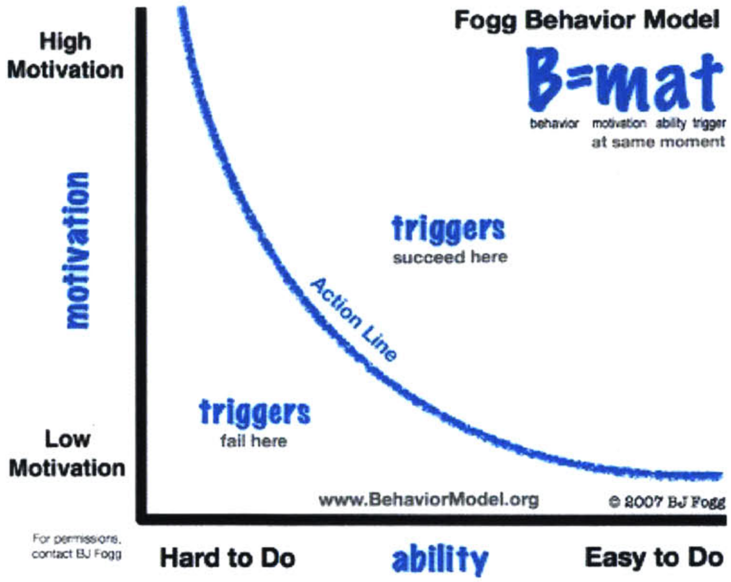


Figure 1-2: BJ Fogg behavior model [1]

1.2 Triggers

As described above, triggers are one of the three variables that make a behavior happen. This thesis investigates the effects of supporting behavior change methodology to wearable technology. Specifically, the thesis investigates ways that triggers can be used as to motivate or re-motivate users. This ability is dependent on the ability to get the user's attention. Even when the system has the user?s attention, timing is key so that the behavior has the best chance of developing. By using technology we can easily fine-tune the content of a trigger so that it is effective. Triggers can take the form of personalized and contextualized messages that are presented just at or before the moment of decision-making. Utilizing technology as a behavior change tool requires designers to consider several factors when deploying triggers. This thesis investigates these different aspects in an attempt to explore how digital systems can be used in behavior change research.

1.2.1 Content

Triggers have the ability to deliver persuasive content as an attempt to ensure the development of a behavior. The content should be able to be quickly and easily digested by the user in a mobile context to prevent further cognitive strain. The content should also be personalized, providing the right information in the right form (factual, emotional, etc.). Finally, the content can be meaningful and/or assist in making the new behavior easier to do. Concise, actionable triggers are ideal.

1.2.2 Motivation

In BJ Fogg's model triggers are described as proactive, used as the final step to create the behavior after optimizing motivation and ability. However from the example we can see that they be used in the moment to change a habit by re-motivating the user and offering an alternative choice. Thankfully it can be assumed that anyone adopting a personal system designed for behavior change is already highly motivated to change their behavior. Therefore, this thesis will explore how triggers can be used to motivate users to perform behaviors that they have not already committed to and how they can make it easier for users to engage in the desired behaviors.

1.2.3 Attention

Attention is often assumed when designing personal systems. However the pervasive nature of technology means there are constantly devices bidding for our attention. In addition these systems are used in a mobile context when they are many other demands upon the user. Therefore because of the just-in-time nature of triggers, personal systems need to be developed in such a way as to ensure that they are able to grab the user's attention so that information can be presented at the opportune moment.

1.2.4 Timing

Another important factor for triggers to be effective is timing. We know it is more effective to influence a user when there are experiencing the situational forces typical for the behavior they are attempting to change[4]. Therefore, it is ideal to give a trigger at, or just prior, the moment the behavior is intended to occur. This thesis will also discuss other aspects that need be considered when timing triggers.

1.3 Thesis Road Map

With the development of sophisticated wearable technology, personals systems offer a means to help combat poor health choices by incorporating behavior change principles. This thesis explores principles for designing applications that help a person be more mindful in their behavior choices. These design principles are investigated through the development of a handful of behavior change systems built with currently available technology. Each of these systems was tested through pilot user studies to learn what the potential challenges are when applying behavior change research in the context of personal wearable systems. The trigger aspects that are being explored are summarized in a table similar to 1.1 for each of the systems in this thesis. Overall, the thesis attempts to develop methods and platforms that can be used to further development of technology that helps users with behavior change.

	Motivation	Attention	Timing	Message Content
Condition 1				
Condition 2				

Table 1.1: Summary of the different aspects of a trigger that is being evaluated in the system

This thesis starts by providing background research on what motivates people to make choices that improve their well-being. The social factors that reinforce these motivations are also addressed. A introduction to persuasive technologies that are health related is then provided. This includes the technological capabilities and designs of current persuasive systems.

The content of triggers messages and the lack of content in triggers are then evaluated. Two systems were developed that allow for the effect of the contents of persuasive triggers to be evaluated. Both systems were evaluated with pilot user studies. Through these studies it is clear that content is important and that the framing of the content can affect behavior change.

This is followed by an investigation into how triggers can be used to increase motivation. Two systems were designed to evaluate the possible relationships between triggers and motivations. This includes how triggers can increase a user motivations and how different motivations effect behavior. Evaluation was done with user studies. The results from these systems provided key lessons for motivating users, mainly the strong influence productivity has on behaviors.

In the attention chapter two pilot studies are presented that evaluate different methods for attracting the user's attention when they are distracted by various environmental factors.

The timing chapter describes a system designed to test how the timing of triggers can effect behavior change. The system is also tested with a pilot user study. The results from the study show that productivity is a strong motivator that should be designed for.

Finally, the lessons learned from designing five systems designed for behavior change are summarized. Other areas of research and improvements to the designed systems are discussed. Overall, this thesis contributes novel systems that address several aspects of well-being. The lessons learned help provide design considerations for future wearable persuasive systems. As this thesis explores a large scope, its intent is not to make concrete conclusions on how best to design systems, but instead find potential challenges when designing such systems and areas that should be investigated further. Using technology as a decision support tool is a fundamental shift, therefore designing systems will require a strong understanding of behavior change research and where it falls short when implemented using wearable technology.

Chapter 2

Background

2.1 Behavior Economics

When designing technology for behavior change it is important to understand the factors that influence an individuals? motivation for behavior change. Motivation is generally categorized as either intrinsic or extrinsic. Where intrinsic motivations refer to actions that are performed because they are self-pleasing or satisfying and extrinsic motivations refer to actions that are performed because they will lead to either a positive or negative consequence. These two types of motivations can influence each other, and together influence a person's behavior. The impact of both intrinsic and extrinsic motivation has been thoroughly explored. Experiments in this area have shown that when a person feels a full sense of choice, they typically exhibit enhanced performance and greater psychological well-being [5]. Therefore the self-determination theory describes the social conditions in which intrinsic and extrinsic motivation produce autonomous behaviors.

2.1.1 Self-Determination Theory

Autonomous motivation, motivation that involves behaving with a full sense of choice, has been shown to result in better job productivity [6] and healthier lifestyles [7]. Therefore it is important to understand how intrinsic and extrinsic motivations contribute to autonomous motivation. Intrinsic motivation is the purest form of autonomous motivation. Generally extrinsic motivation only works to decrease intrinsic motivation and inherently a person's feeling of autonomy. Previous research has shown that adding extrinsic rewards, such as prizes or money, almost always leads to a decline in intrinsic motivation [8]. However, people can feel autonomous while being extrinsically motivated. This occurs when extrinsic motivation is internalized. Factors that effect internalization of extrinsic motivations and the feeling of autonomy include positive feedback and goal setting.

2.1.1.1 Positive Feedback

Positive feedback, while extrinsic, has been shown to increase intrinsic motivation [9]. Positive feedback is also a way to increase self-efficacy, or a person's ability to believe they are capable to succeed at accomplishing a specific task. Self-efficacy is a predictor of short and long-term success and a powerful factor in initiating and maintaining health behavior change [10]. Therefore, it is a natural conclusion that technology can be used to increase a users intrinsic motivation and feeling of autonomy by providing positive feedback to measured actions that are aligned with his/her behavior change goals.

2.1.1.2 Goal Setting

A study of goals and their effect on well-being found that goals fall into two categories, intrinsic and extrinsic. Intrinsic goals were shown to lead to higher well-being and less stress [11]. This result was further strengthened by findings that showed the negative relationship between extrinsic goals and well-being was stronger than the positive relationship between intrinsic goals and well-being [12]. This shows that trying to motivate a user with monetary rewards is more detrimental than allowing a user to create goals that align with their personal goals. Similarly manipulating a person's goals by making the goal either intrinsic or extrinsic in natural can effect how well material is retained [13]. However, setting a stagnate goal or an unachievable goal, regardless if it is intrinsic, has negative effects. Goals should vary with performance, provide a path to a larger goal and be challenging [14]. Once again, it is clear to see how technology can be used to set intrinsically motivating goals that are personalized to the user.

2.2 Persuasive Technology

Persuasive technology, technology designed to change attitudes or behaviors of users through persuasion and social influence, can be implemented for a wide variety of behaviors [15]. Persuasion has long been a tactic used in marketing and advertising. These same tactics were deployed on websites with the rise of digital information. However, as we move to wearable technology, it becomes easier to use aspects of above described self-determination theory. As this thesis focuses on behavior change for well-being persuasive technology in the well-being space is only discussed. Persuasive technology tends to have three stages: measurement, persuasion and feedback.

2.2.1 Measurement

Current technology has made it possible to include a variety of sensors in one device. This makes it possible to detect a wide range of user actions and characteristics. GPS has made it possible to know where a user is at all times and is available in most smartphones. Location can be pinned down further with the use of Bluetooth and RFID's [16]. What a user is doing in a location can be determined using accelerometer data. One study found smartphones able to detect step count with a relative difference in mean step count ranging from -6.7% to 6.2%. When wearables are used the accuracy ranges from -22.7% to -1.5% [17]. Social activity and emotional state can even be measured through the use of wearable technology. Electrocardiogram, electromyogram, skin conductance, and respiration have been used to measure stress with an accuracy of 97.4% [18]. Through the use of computer vision, technology can also detect faces and accurately determine their expression [19]. Coupled with speech recognition a more accurate detection of emotion can be created [20]. With these sensors and techniques, as well as others not described, an accurate picture of the user's current state can be created.

2.2.2 Persuasion

Once current state is measured, persuasion techniques can be applied to either reinforce behaviors or trigger new behaviors. The variability of software makes it possible for persuasive technology to utilize a variety of tactics when persuading a user. The most popular tactics include social influence, personalized goals, and education. With social influence, the user is made aware of their performance in comparison to others. This generally creates a competitive atmosphere and helps to set achievable but challenging goals [21]. Personalization is used to provide an experience that is specific to the user. Systems that have utilized personalized messages for their users have shown better results [22]. Finally, education is used to aid in a changing the user's attitude. If the user accepts the importance of the desired behavior as a result of the new information, the feeling of autonomy and compliance can be increased [23].

2.2.3 Feedback

Performance feedback is typically provided through a variety of methods. The most common method is visualization. Data trends are displayed to the user to easily illustrate change over time. This is particularly common for weight loss related goals. Recommendations are another tactic. Users are given messages periodically to suggest actions for improving performance. Finally, rankings are provided to compare between the user and their community. As stated above, this helps to encourage competition and effective goal setting.

2.3 Research in Behavior Change with Wearables

In recent years, research has been done that builds on the above mentioned behavior change theories and adapts them to be used passively with wearable technology. Current research is investigating methods to accurately detect behaviors and activities, set appropriate goals and deliver just-in-time messaging, as just a few examples or research being done in this space.

2.3.1 Detecting Activities

As persuasive systems are designed, it becomes neccessary to detect the current user's state to provide contextual feedback. Research in this area has studied wearable sensors and sensors placed within the environment. For example, wearing five accelerometers allows a researcher to detect everyday activities such as watching TV, vacuuming, and stretching [24]. Environmentally based sensors are location dependent, but do not require precious real estate on the body. These sensors can also detect everyday activities, including washing hands and grooming [25]. While wearable sensors can help to accurately detect physical activities, environmental sensors can help detect location dependent trends.

2.3.2 Goal Setting

Goal setting can be effective when goals are intrinsic in nature. Research is currently being done to determine how to best assist users with goal setting based on their personal desires. Research in this area has shown that when using mobile technology to encourage physical activity, users benefit from being able to concurrently pursue primary and secondary goals every week. This allows users to have a better chance of achieving at least one of their goals. If a user is doing especially well they can push to achieve both goals [26]. Similar research in this area has found that weekly goals are preferred. This allows users to reset their goals on a week-by-week basis. Finally, users prefer to set their own goals rather than following national guidelines or goals set by a fitness expert [27].

2.3.3 Just-in-Time Messages

Just-in-time messages can be a viable way to motivate behavior change with real-time feedback or suggestions to help with goal completion. Research has been conducted to determine what types of messages users are most likely to respond to when presented by a personal device. Messages were found most credible when they came from a health expert. Additionally, studies have concluded that messages should not include negative enforcement or complex suggestions [28].

2.4 Current Persuasive Systems

By applying behavior change research and lessons learned from applying this research to wearable technology, new persuasive systems have been developed. Systems in the persuasive computing space can be divided into three categories: systems that augment the environment, systems that require wearables or technology solely for the intended behavior change and systems that do not require new hardware but instead take advantage of devices and technology the user already utilizes for other purposes. Most systems attempt to address one or all of the three variables in Fogg's behavior change model, namely increasing motivation, increasing ability or offering a trigger.

2.4.1 Augmenting the Environment

Augmentation systems are persuasive systems that tend to augment everyday environments by adding technological features. These systems are generally limited to a specific location and if capable of sensing, only gather data from that specific environment. However, since these systems augment objects that are already naturally used, they do not have to overcome adoption hurdles. Augmenting objects in a way that they can still be used naturally and sense or display data is not easy though. There have not been many that have been made commercially available. This may be because of the high standard that they are given. For users to not resort to their original non-augmented forms, the new augmented version most be accurate and reliable and unobtrusive.

Systems in this space include EducaTableware [29], Darma [30], Waterbot [31] and Mug-Tree [32]. Each of these systems attempt to further motivate a user to perform a behavior they are already committed to but struggle to do on a regular basis or want to do better. However, the method each system uses varies and their dependence on a location makes them unlikely for long-term adoption.

EducaTableware attempts to make eating more pleasurable and and encourage daily eating habits for kids who resist eating. The system includes a fork and cup that emit sounds when used. This creates a habit loop where eating and drinking becomes associated with a reward of auditory feedback. While the novelty of the system may encourage eating for kids initially, the predictable response can result in the feedback system to become boring or even unpleasant, ending the habit loop.

Darma, an augmentation for the chair, measures posture and sitting time and provides exercises and feedback to the user through a smartphone application. If a user sits for too long or maintains bad posture, they get a trigger from their smartphone, which reminds them to change this behavior; they are then rewarded for their action through the app. The use of an app provides the opportunity for variable content in the feedback process. The ability to give users messages in the moment of their action increases the chances of behavior change. However, the device is rather large (to cover the complete surface of a chair) and must be removed and re-applied to each chair a user intends to use.

Mug-Tree enforces adequate water consumption through a mug and digital photo frame. As users drink from the cup (by tilting it) the event is transmitted to a photo frame that visualizes the amount of water consumption with a tree that is either flourishing or dying based on the user's water consumption. The act of drinking is rewarded by the visualization of a healthy growing tree. This could be further enforced socially when others notice a user's tree and associate this with the users drinking habits.

Finally, Waterbot is a persuasive system that encourages water conservation by visualizing water consumption on a sink faucet. The faucet provides auditory feedback when it is turned off and visual feedback shows the current user in comparison to others. Users are triggered to reduce their water consumption with these feedback methods and rewarded by seeing their consumption as compared to others in their household.

Each of these systems works to motivate and reinforce positive behavior. However, since they are built for a specific tool and specific environment, the habit loop they create tied to the use of the device. Carrying these devices around so that they can be used in every setting is unlikely. The feedback method is often limited in variability. As users become accustomed to the feedback, the positive reinforcement may wear off before the habit is formed. While augmented persuasive systems are highly effective at motivation and offer ideal timing, their dependence on location makes them not ideal for pervasive behavior change.

2.4.2 Wearables

Several persuasive systems attempt to combat the limitation of augmented environments by designing wearable systems intended for a specific behavior change. The most popular of these is the ever-growing list of fitness trackers (Fitbit [33], Jawbone up [34], Misfit Shine [35], etc.). These devices monitor a user's activity level and provide feedback through visual data and messages via a smartphone app or haptic response. Other novel wearable systems include Fit4Life [36] and Spire [37].

Systems in the fitness tracker category are typically wrist worn devices that have evolved from the popularity of pedometers. These trackers at a minimum track steps and can differentiate between states such as walking, sitting and running. Typically trackers also offer other features such as sleep tracking, heart rate monitoring, goal setting and tailored feedback. For most users wearing the device becomes a trigger to be more active and their movement is rewarded with positive reinforcement through tailored messages, visual and/or auditory feedback. Most devices are designed such that they don't need to be taken off (except to charge); therefore users are more likely to utilize the device for long-term behavior change.

The Fit4Life system is a multi-device system that tracks user's activity, provides feedback and incorporates social support. The system includes a heart rate monitor to measure activity levels, an earpiece to provide auditory feedback and reinforcement, an app to collect data and a beacon accessory to elicit help from others. Data sent to the cloud support portion of the system can post messages to social media to further elicit encouragement and celebrate successes. There have been several systems similar to this that utilize several devices so that accurate data can be collected from multiple locations on a user's body.

Finally Spire, a wearable device that monitors respiration, represents a growing trend in wearables to track more than just activity. Spire attempts to help users be more mindful and relaxed by controlling their breath. The accompaning smartphone app guides users through relaxation techniques to reduce tension and lower blood pressure.

Each of these novel wearable systems is effective at giving users information about their current activity levels. For some people this leads to short-term behavior change. The portability makes them a more viable attempt at long-term behavior change as users are able to take their trigger with them to be used in a variety of situations and environments [38]. The biggest barrier these devices face is slow adoption. As wearables become more ubiquitous, users will not want to adopt yet another device that provides a single function. Users that are not quick to adopt technology will also be slow to take advantage of these systems. Systems that include multiple devices, such as Fit4Life will be the last to be adopted, as it requires large amounts of real estate on the human body. Therefore while novel wearables are efficient at motivation and habit formation, there adoption is limited because they are generally only serve one purpose. The majority of users do not utilize these wearables in the long term. 32% of users stop wearing their device after six months and 50% after one year [39].

2.4.3 Utilizing Existing Technology

Systems that utilize existing hardware technology are the main focus of this thesis. Systems in this category take advantage of existing technology and add new capabilities. By using technology the user has already adopted, the probability of endorsement is higher. As smart phones have become more widely used, a variety of persuasive applications have been designed to collect data from sensors and give feedback through alerts. Similarly, new devices such as smart watches have included features similar to fitness trackers discussed in the section above.

A plethora of smartphone applications have been designed to encourage user behavior change. These applications include those designed for nutrition tracking (MyFitnessPal, Lost It!, MyPlate), sleep tracking (Sleep Cycle, Sleep Better, Sleep++), activity monitoring (Human, Runkeeper, MapMyFitness), meditation guidance and tracking (Calm, Headspace, Stop Breathe Think) and general goal tracking (Way of Life, Productive, Coach.me). Each of these applications utilize ssensors on the phone or input from the user to gather data and trigger a behavior. Points typically reward adherence and trends are shown over time. Adoption is very easy as a user typically only need to download the app and most basic versions are free.

Other systems take advantage of more novel wearables such as the Apple Watch or Google Glass. While users typically use these devices because of their promise of increased productivity and connectivity, designers are able to offer the added benefit of applications designed to improve well-being. The Apple Watch, shipped with applications that aid the user in well-being related behavior change such as providing a haptic response every hour to remind the user to get up and take a break.

Systems in this category are able to develop and deploy to a user quickly. They take advantage of technological advances from existing technology that results in a much shorter development period. However, since they are designed for existing devices, they are limited by their sensing capabilities and rely on additional data from the user to be fully useful. Additionally, the ease of development for systems allow designers to quickly create systems regardless of if they incorporate behavior change principles and design considerations. Developers are typically more interested in short term adoption rather than long-term effectiveness.

2.4.4 Summary of Existing Systems

Table 2.1 summarizes the pros and cons of each of the system categories discussed above. While each category of system has its weaknesses, the weaknesses that exist in systems that utilize existing technology are easy to overcome. More thoughtful design of these systems can overcome the need for human input and make them more effective for long-term behavior change. Additionally, with the rapid improvement of smartphone sensing capabilities and introduction of newer wearable devices, the accuracy problem that currently exists is temporary.

Category	Pros	Cons
Augmented Environmentst	• Provide just-in-time feedback • Ensure they have the attention of the user • Adoption barrier is low as items are typically familiar to user already	• Limited to a specific location and environment • Often do not provide variability in feedback responses • Challenging to design systems that are seamless and accurate
Novel Wearables	• Provide just-in-time feedback • Ensure they have the attention of the user • Able to be used in a variety of locations and environments	• Adoption barrier is high as systems are typically singular in focus and unfamiliar to users • Systems that monitor a variety behaviors require more human real estate.
Utilize Existing Technologies	• Provide just-in-time feedback • Ensure they have the attention of the user • Adoption barrier is low as items are typically familiar to user already • Quick development time	• Often require additional input from user to be effective • Typically designed without taking into account behavior change principles

Table 2.1: Pros and cons for existing categories of persuasive system

Chapter 3

Experimenting with Message Content for Behavior Change

3.1 Introduction

Self-monitoring behavior can be seen in behavior psychology as early as 1970 [40], with monitoring specifically for weight loss appearing in the literature as early as 1985 [41]. These behaviors have grown into what can now be called the quantified-self movement, with people diligently tracking data about themselves through the use of technology [3]. In a nation-wide study it was found that 69% of U.S. adults track a health indicator like weight, diet, exercise routine, or symptom. Of those, half track "in their heads," one-third keep notes on paper, and one in five use technology to keep tabs on their health status [42]. The number of technology trackers will surely rise as we develop devices that accurately and seamlessly collect data.

Tracking data provides insight into behavior and opportunities for change. However, according on the Fogg Behavior Model [1], change in behavior may not occur without a reliable trigger for the new behavior. Tracking also does not provide aid at the moment of decision-making. It may be more effective to influence a user when they are experiencing the situational forces typical for the behavior they are attempting to change [4]. Wearable devices have an opportunity to do so at the right moment. While timing is important, the content of the messages presented to the user can have an impact on the message's effectiveness. A lot is known about the effectiveness of messaging in marketing efforts [43], however less is known in regards to the effectiveness of these marketing techniques in the context of in-the-moment decision-making. Typically messages that are personal to the user (fit), arouse an emotional response or provoke a competitive atmosphere (gamification) are the most effective in health related marketing [44]. This chapter first investigates whether messages are necessary for a specific behavior change or can the knowledge of a user's activity patterns be sufficient to persuade them to change their behavior. It them describes a system and a user study developed to test which messages are most effective when used as a just-in-time trigger for a specific behavior. In this chapter two systems are described. The first attempts to effect snack choice, while the second attempts to impact time management.

3.2 Food Attack

Making healthy food choices is difficult for many people. While it is easy to set diet goals, it is hard to keep them when making in-the-moment food choices. The Food Attack system utilizes a head mounted display to present health-based messages to users as they make real time snack choices.

The game is meant to prove that just-in-time messaging is a viable approach for behavior change. As shown in table 3.1, the Food Attack game utilizes each aspect of a structured trigger. Persuasive messages are presented in-the-moment in a way that grabs the user's attention.

Motivation	Attention	Timing	Message Content
√	√	√	√

Table 3.1: Summary of trigger usage for Food Attack system

3.2.1 Related Work

A 2012 study found that the iTunes app store contained 1,336 apps in its Health & Fitness Category [45]. A large number of these apps focus on food choices or logging food choices. However, all of these apps require input from the user and for the user to utilize the app as intended. They lack the ability to present messages to users at the moment of decision-making and are typically not backed by behavior change theory [45].

There have also been attempts to design technology that passively detects eating habits and persuades in real time. However, many of these designs focus on children [46][47] and require the addition of a new device that is only used for eating [8]. These devices also focus on the rate of food consumption rather than food choices, as there are limitations in automatically detecting food calories and nutritional information.

Psychologists have done substantial research into the most effective messages for health related marketing [44]. However, this work usually focuses on one-time decisions and is difficult to deliver at the time of decision anywhere, anytime. These messages are also not personalized and are location dependent. The Food Attack system is designed for wearable technology that is not restricted by location. By using a head mounted display, there is a higher probability that the user views messages wherever they are and whatever they are doing. It also allows for messages to be personalized so they are aligned to the user's desires, their unique personality or their current situation.

Food Attack shows that a simple just-in-time solution to aid in food choices can be built using existing head mounted displays already adopted by some users. It does not require the detection of food choices, but sends triggers at opportune moments to encourage healthier choices.

3.2.2 System Design

A game designed for Google Glass as a simple testing platform for just-in-time messaging. Each game lasts no longer than 20 minutes. During the game the user is tasked with collecting as many objects as possible before the end of the round. Objects appear at random locations on the google glass screen for 5 seconds. Participants can control a crosshair on the screen by moving their head. Once a crosshair is placed over an object, they can execute a tap gesture to "collect" the item. Each round of the game lasts approximately 2 minutes followed by a 2-minute break. During the break the participant was shown a message and a timer counting down till the next round. The break provided "dead time" that was ideal for snacking as participants waited for the next round of the game to begin.

After the timer concluded (during the dead time), the user would perform a tap gesture to start the next round. The game ended when the participant completed 5 rounds. Prior to starting the game, users were given an introduction to Google Glass and the game, which walks them through executing tap gestures, moving the crosshairs and adjusting the Google Glass to their comfort.

Figure 3-1: Screen shot of Food Attack.

The game could be played in 4 different modes. The different modes were used to determine which variable in the game was responsible for any influences. The control mode (mode 1) of the game did not show a score to the participant and used different balls as the objects to collect. After each round the participant was presented with a message stating they were doing a great job and asking them to wait for the next round. Mode 2 replaced the objects with different food items. During this mode participants were still not shown a score. In mode 3 participants were given a score based on the food objects they collected. Collecting healthier food items resulted in an increase in score, while unhealthy food items resulted in a decrease in score.

Participants playing in this mode were given a key showing which items would be displayed and their point value. In mode 4 participants were given one of 3 types (emotional, gamified and fit) of persuasive messages during the break period and shown a score. Each message was meant to encourage eating healthy foods and discourage eating unhealthy foods. There was a unique message for each round of the game that was played. Table 3.2 summarizes the 4 modes and includes an example message for the different message types (the full list of persuasive messages used can be found in appendix A.1). Emotion based messages focused on hedonic messages that described a direct correlation between food choices and a person's health. Fit based messages were also hedonic, but focused on health aspects that are generally a priority to the user based on gender. Finally, gamified messages focused on how the user could improve their score in the game by choosing the healthier food items.

Mode	Objects Used	Score Used?	Example Message
1	Balls	No	Wow, you are doing great! We are getting the next round ready.
2	Food	No	Wow, you are doing great! We are getting the next round ready.
3	Food	Yes	Wow, you are doing great! We are getting the next round ready.
4	Food	Yes	**Emotional:** Eating healthy foods makes you feel good. **Gamified:** Eating fruits may help improve your score. **Fit (Men):** Vegetables can help with building a lean body. **Fit (Women):** Eating healthy foods makes your waistline smaller.

Table 3.2: Food Attack game modes

3.2.3 Study Design

Participants for the study were recruited with an invitation to try a Google Glass game. Participants were not told the intention of the study, but did know that snacks would be available. Overall 90 people participated in the study. The distribution of participants can be seen in Table 3.3. Each participant completed the study in a cubicle with no other participants immediately to their right or left. Prior to being seated a bowl of snacks was placed at the desk. Each bowl contained exactly 10 M&M's, 10 grapes, 10 skittles and 10 baby carrots. Once shown their testing area, the participant was given an introduction to Google Glass that included the basic mechanics. Before they were able to begin the game they were fitted with Glass and made sure they were comfortable and could adequately see what was presented on the screen. At this point they could start the introduction of the game and continue the study. Each study session had at most 5 participants. Participants were not able to see other participants' testing areas to avoid social influence in regards to food consumption.

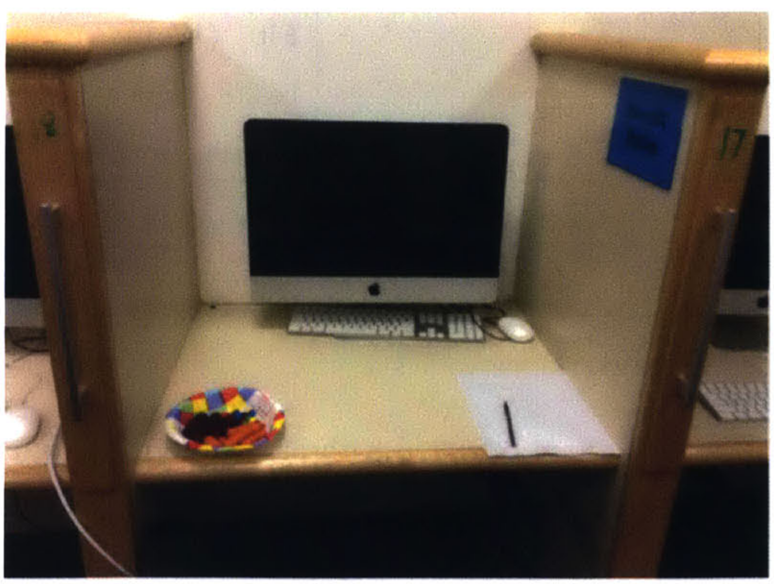


Figure 3-2: Participant testing area

Mode	Percentage of Healthy Eaters	Number of Participants
1	13%	15
2	29%	14
3	54%*	13
4. E	56%*	16
4, G	33%	15
4, F	29%	17(8 Men)

Table 3.3: Percentage of healthy eaters by game mode. Significance designated with * ($p<.05$).

3.2.3.1 Study Analysis

When the participant received the "game over" screen they were instructed to complete the post-survey left face down in their testing area. The survey asked about the timing of their last meal, stress level of the day, feelings about their performance, snack consumption and feedback for the game. After a participant completed the study the remaining food items were counted. Grapes and carrots are classified as healthy items while M&M's and Skittles are classified as unhealthy.

3.2.4 Results

To determine if any of the modes were effective at persuading the participants to make healthier food choices, the number of people in each mode that choose to eat healthy was analyzed. Eating healthy was defined as a person that consumed more healthy snacks than unhealthy snacks (consumption of grapes and Skittles >consumption of M&M's and Skittles). These results are summarized in Table 3.3. Mode 3 and the emotional subset of mode 4 had significantly more people that chose to eat healthier as compared to the control group. There was no mode or subset that consumed significantly more snacks as compared to the control group, as seen in Figure 3-3. This further shows that participants in mode 3 and the emotional subset of mode 4 chose to make healthier decisions when snacking during the study.

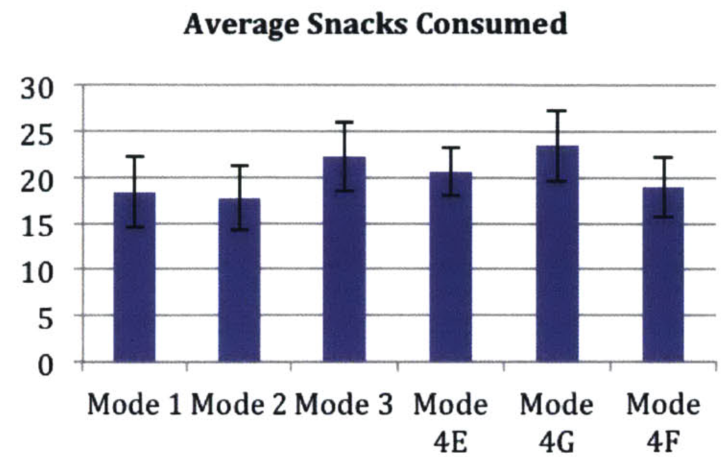


Figure 3-3: Average number of snacks consumed per participant during the study by each mode and subset

During the post survey, participants were asked to rate how they felt about their snack consumption on a 7-point likert scale. Those that were in the emotional subset of mode 4 rated somewhat significantly higher, in comparison to the control group, how they felt about their snack consumptions ($p = .066$). This may show that those in this category felt better about what they chose to snack on during the study. These results can be seen in Figure 3-4.

Participants were also asked to rate how they felt about their reaction times on a 7-point likert scale. The mean scores for each mode and subset can be seen in Figure 3-5. Those in the emotional subset of mode 4 significantly rated themselves lower than the control group ($p=.005$). This strongly suggests that after playing the game with emotional messages, participants felt worse about their reaction time and performance in the game. A possible explanation is that participants compensated for their perceived poor performance by choosing healthier snacks to improve their emotional state.

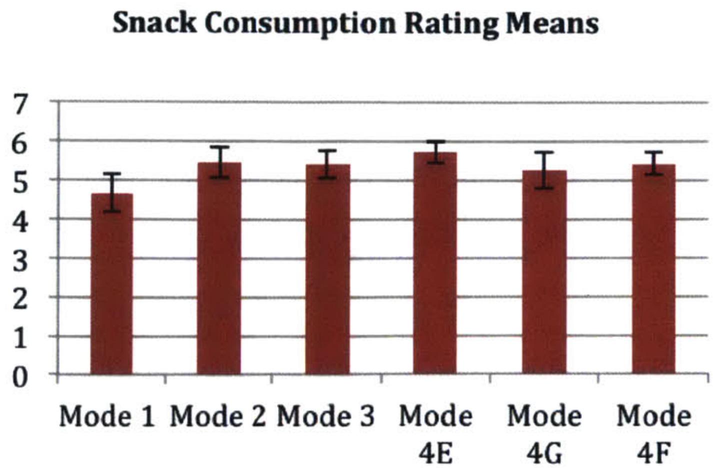


Figure 3-4: Average healthy snack consumption by mode

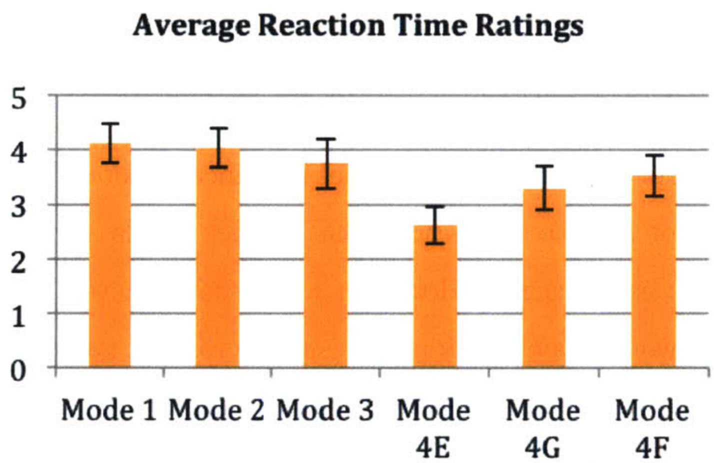


Figure 3-5: Average reaction time response by mode

3.2.5 Discussion

Food consumption can be highly influenced by previous meals and time in-between meals. To attempt to factor out these conditions, each participant was asked when their last meal took place. These results are shown in Figure 6. While those in mode 3 had on average less time since their last meal, the average time, as compared to the control group, was not significantly different. This is confirmed when the participant's food score (number of healthy foods chosen) is graphed against time since last meal as seen in Figure 3-7. There is no clear trend between a participant's last meal and the snacks they chose. This is also true of the number of snacks eaten and the time since last meal as seen in Figure 3-8. Each of the modes was performed at different times during the day. This strongly suggests that neither the time since a participant's last meal nor the time of day influenced the amount or choice of snacks participants consumed.

While previous studies find that fit based messages are most effective for health communications [44], we found emotion based messages to be more impactful. The emotional connotations with food and body image may make these messages more effective for users attempting to make healthy food choices (mode 3). We also saw that simply priming users to choose healthier foods through the use of scoring mechanism can impact food choice. However, those that received emotional based messages experienced the same effect with the addition of improved reaction to their choice making this the preferred method.

An effort was made to minimize any social influence on food choice. Although the participants were not able to view each other's snack bowls, it was audible when a participant chose to eat a carrot. However, we do not believe this influenced other participants or we would have significant results across all game modes since there was someone in each session that chose to consume carrots.

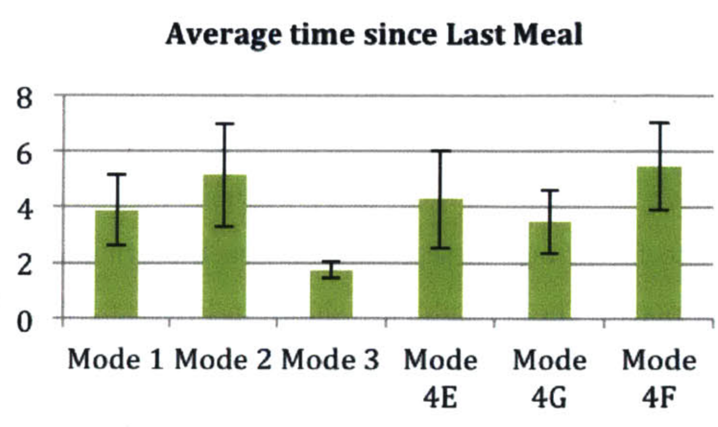


Figure 3-6: Average time since last meal in hours for each mode and subset

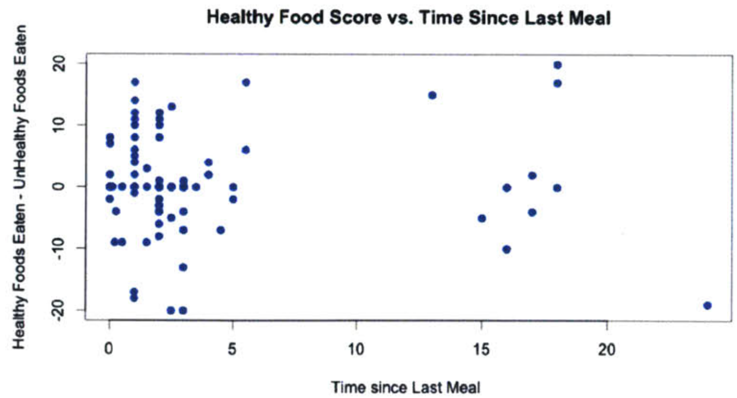


Figure 3-7: Number of healthy snacks consumed vs. time since last meal

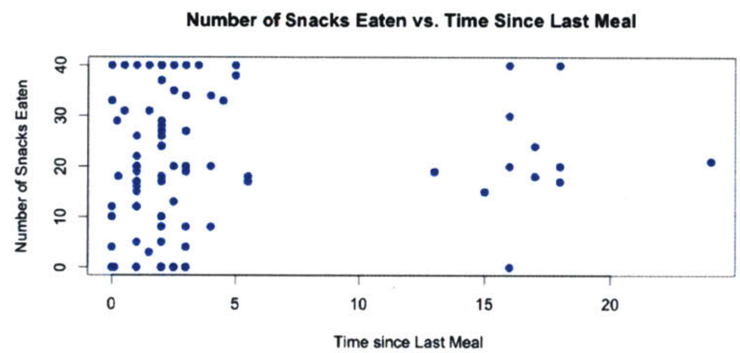


Figure 3-8: Snacks eaten vs time since last meal

3.2.6 Study Conclusion

The Food Attack experiment helped show that giving users persuasive messages at the time of decision-making is an effective way to encourage healthy eating choices. While other solutions focus on introducing new technology or accurately detecting food choices, this solution can be incorporated in current technology and only needs to detect when a user is eating. Sensing food consumption can be passively done in a variety of ways [48]. The above study shows evidence for what types of messages are most effective for users. Messages can be changed periodically to keep them interesting to users and unpredictable. The rise of wearable devices has made it possible to present messages quickly and even at eye-level, increasing the chances for these messages to be seen and processed. Exploration could also be done into the use of imagery instead of text-based messages. While this study shows effectiveness at the moment of consumption, similar principles could be applied to the design of human-computer interactions that involve messaging at the moment of meal selection at restaurants, vending machines, etc. Finally, this study shows evidence that the content of the message can have an impact in the effectiveness on the resulting behavior. Therefore, it is important to spend time designing messages that are both interesting and effective. The lack of messaging is explored with the Watch system.

3.3 Watch

Productivity is a strong motivator for most users. However, time-management and efficient use of time continue to be skills that are hard to master. Additionally, users feel that they are subjected to increasing time pressures and overall pace of life [49]. It is therefore no surprise that we increasingly rely on technology to keep us on track and help plan the most effective use of our time. WATCH is a time-management tool that we developed that visualizes how a user has spent their time in comparison to their goals. Currently, the system exists as an android app that utilizes the Jawbone API to track users' activities. The two-week study designed to test WATCH variation in timing of the message, however both weeks do not utilize messaging as seen in table3.4.

	Motivation	Attention	Timing	Message Content
Week 1	√	√	X	X
Week 2	√	√	√	X

Table 3.4: Summary of trigger usage for Watch system

3.3.1 Previous Work

There are several apps designed to allow users to track how they spend their time [50]. These apps rely on user input to classify the activity and typically include a timer to track the time spent on a pre-determined activity. While some allow users to set goals or the maximum time they would like to spend on an activity, they all focus on tracking one activity at a time and only support time spent on mental work, therefore they are only useful for a subset of the entire day. Users are forced to use more than one tracker to effectively track other aspects of their time. The Watch system is able to track time spent in a variety of activities and gives a more holistic view of how time was spent during a day.

Other systems have attempted to track time passively with the use of wearables and sensors available on most mobile devices. The Life Logging system [51] also utilizes a wearable activity tracker to monitor activity along with sensors in the phone. With the data it collects it attempts to quantify a user's quality of life based on four indicators: activities, sleep, fatigue and mood. The overall scores and how time was spent is presented to the user as a weekly total. While useful information, it is hard for a user of the Life Logging system to turn this data into actionable steps that they can perform that day to modify their behavior and thereby their use of time. Instead of relying on an additional device, the automatic life logging system used the microphone on a user's mobile device to listen and detect activities [52]. This data was used to make a virtual model of a user's daily life in terms of where they spend their time. These systems do not allow the user to set goals or witness in real time what activities they have performed. However, monitoring time based on where a user spends his/her time is not common in other time-management systems. The Watch system allows the user to see their current activity as compared to their goals in real time. This allows users to quickly see how they can change their actions to more closely match their goals. The system also allows users to visualize their time in terms of activities and locations.

3.3.2 System Design

The Watch system has three components, an app to view the visual data and set location parameters, a widget to display the data on the home screen of a user's phone and a website to link the user's jawbone account, and set time-management goals.

Time management goals are divided into two categories, location based goals and activity based goals. Each category has three possible values and accounts for 24 hours. Time in the locations based category is tracked as either time spent at home, work or other. Time in the activity-based category is tracked as active, inactive (i.e. sitting) or sleeping. Users are able to set goals for the amount of active and inactive time, as well as the total time spent at home and at work in a 24 hour day. These goals are compared to the actual measured data for each time category. This comparison is presented to the user with a double donut style pie chart on the phone's homescreen as seen in Figure 3-9.

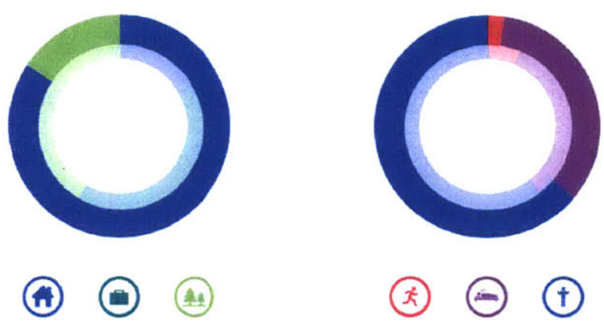


Figure 3-9: Screen shot of Watch App. The inner ring signifies the users goal, while the outer ring signifies the measured data. The 6 icons represent, in order, time spent at home, time spent at work, time spent at other locations, active time, time spent sleeping and time spent being inactive.

By using the app users are able to see data visualization and mark their current location as either home or work using their GPS data. Marking home and work locations are only done once, the app then passively tracks the users location. Through the website, users can authorize the Watch system to use data collected through a Jawbone activity tracker. Users also have the ability to set their goals for the above mentioned category sections. The widget, which lives on the home screen, shows a smaller version of the data visualization, so that uses can quickly see how they have spent their day so far and how their day so far compares to the their time-management goals every time they use their phone. The widget serves as a constant reminder of the user's goals and where they are falling short in meeting them. Users also only see these visualizations during breaks when they are using their phone, so the app does not compete with productivity goals. The system was tested with a user study to determine if quantified data about personal behavior and optimized timing alone is effective at inducing behavior change.

3.3.3 Pilot Study Design

To test if the system was effective at persuading users to meet their time management goals, a two-week study was conducted with 31 participants (15 women and 16 men). Participants were recruited by advertising for users interested in testing a time management application. The only requirement for participation was owning an Android device, being willing to download two applications to their phone and wearing an activity tracker for the duration of the study.

Each participant attended a pre and post study session. During the pre study session they completed a pre-study questionnaire, received a tracker and walk-through of the application. Before they left, each participant was shown how to work the Up tracker and the apps and set their goals. During the two weeks each participant was expected to wear the tracker at all times and track his or her sleep by turning the tracker into sleep mode (and turning it back to active mode in the morning). At the post study session, participants returned their trackers and completed a post-study survey.

In the first week of the study participants were not able to view their data as this was treated as a control week. The widget stated that there was "no data to display". At the start of participant's second week, they were taken out of control and able to view their graphs as seen in Figure 3-10 and Figure 3-11. The outer ring reflects the measured data with GPS data from their mobile devices and data from the Up activity tracker. The inner ring reflects the participant's goals. It is therefore easy to make a comparison between goal and actual data and see the discrepancies.

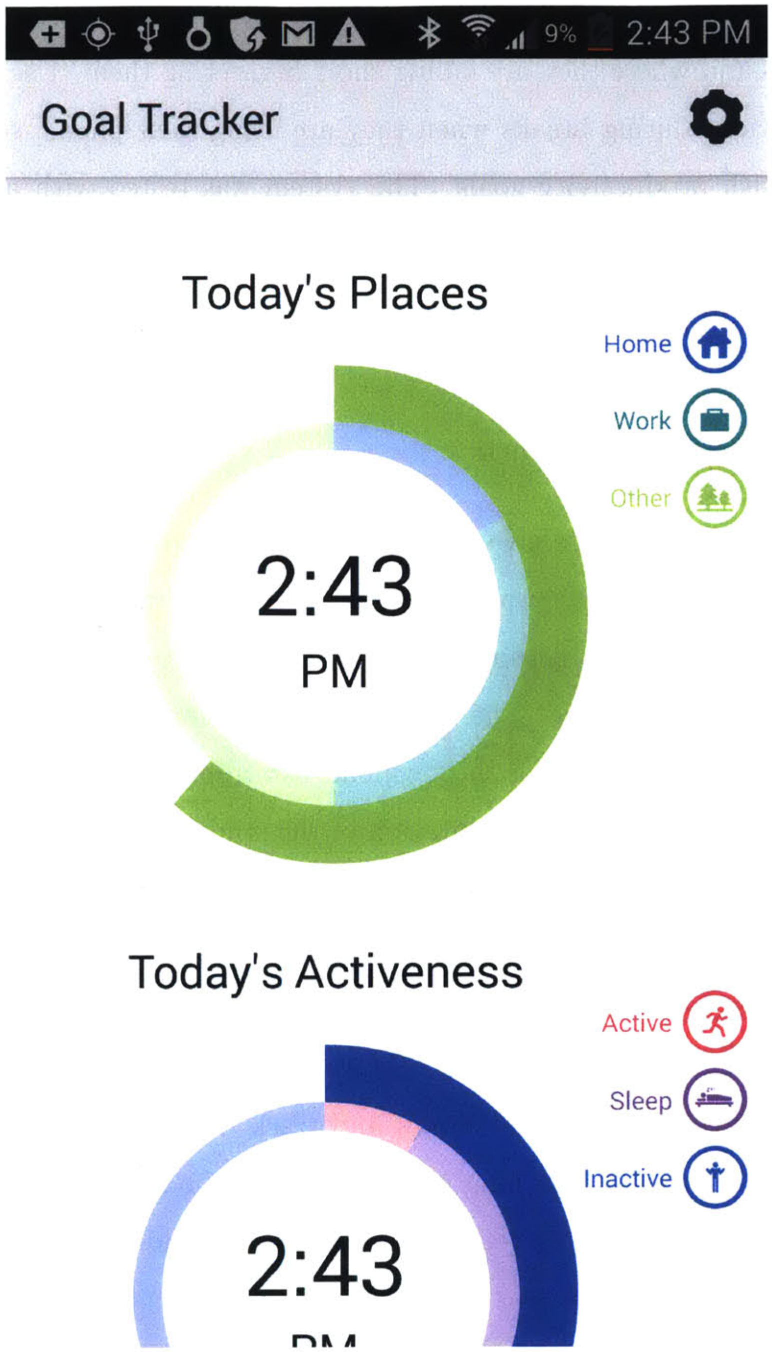


Figure 3-10: Screen shot of real time data in the Watch app.

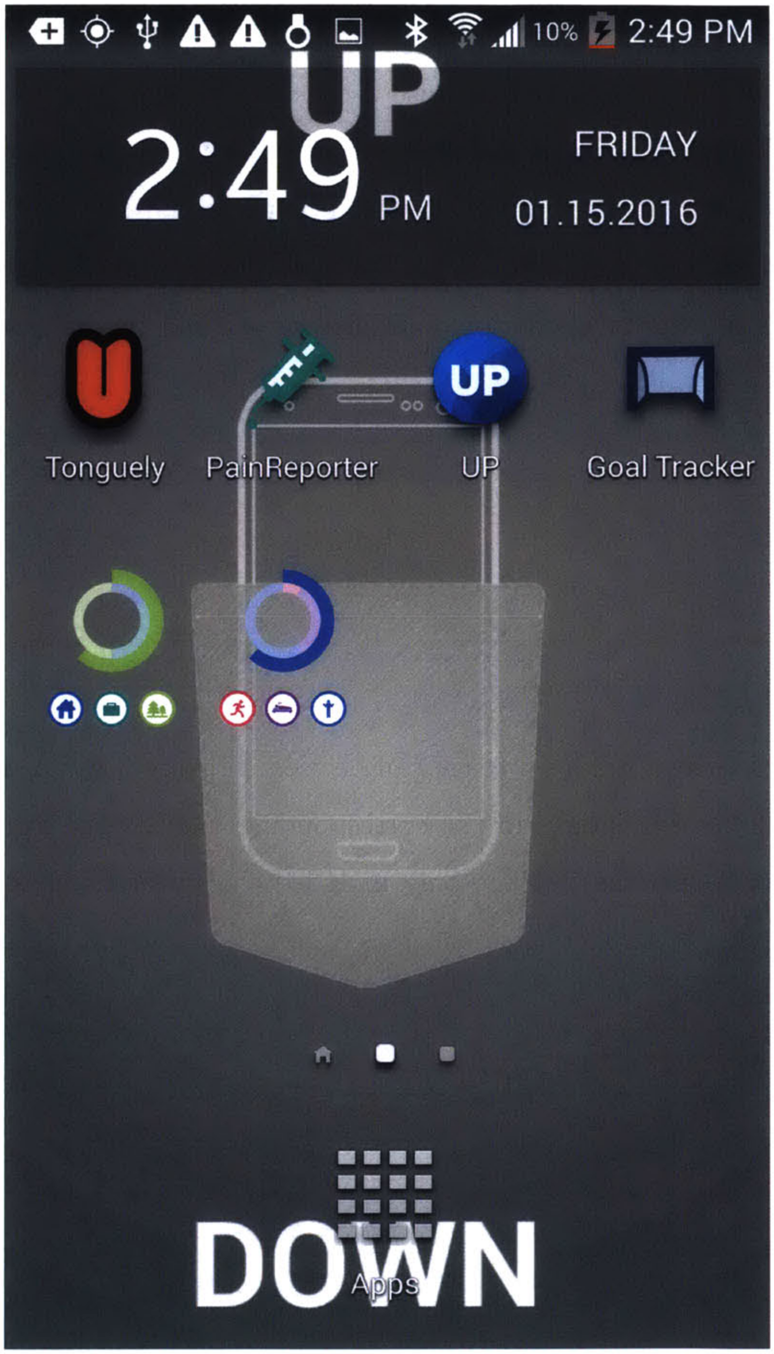


Figure 3-11: Screen shot of Watch widget.

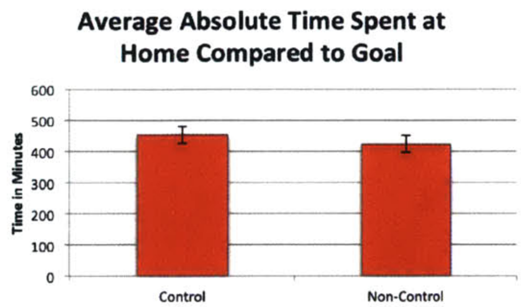
3.3.3.1 Results

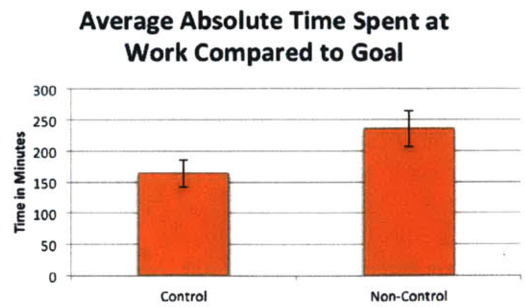
To evaluate the effectiveness of the Watch App, data from the user's control week (week 1) was compared to the second week. To determine if users improved their daily distribution of time as compared to their goals, five metrics were evaluated:

sleep, sedentary time, active time, time at work and time at home. Time spent in "other" places was not evaluated because users did not set a goal for this qualifier. The absolute difference between the goal and the collected time was evaluated first. As many of the participants did not have traditional 9-5 jobs in a set location, they were excluded from the work evaluation. Many of those excluded were students who did not have a set working location, but instead studied in their rooms/home or several places on campus.

The results for the absolute average can be seen if Figure 3-12. A user's data was only included if there was data in the corresponding metric. For example, a user's activity data for a specific day was only included if they had a non-zero value for active time (having 0 activity time suggests a user didn't wear the tracker as basic movements would result in some active time). The same logic was applied to time spent at home and work.

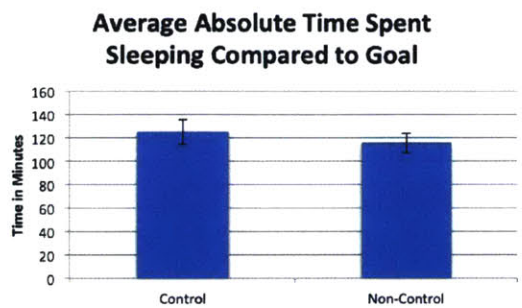
While the data in Figure 3-12 shows how close users were to their goal, it is also interesting to note the distribution of data around a user's goals. Figures 3-13 and 3-14 show the distribution of measured data from week 1 and 2 as compared to the user's goals (denoted with a '+'). It is interesting to see that most users either consistently overshot or undershot their goals suggesting that they are unaware of how they currently spend their time and what would be reasonably goals to aim for.

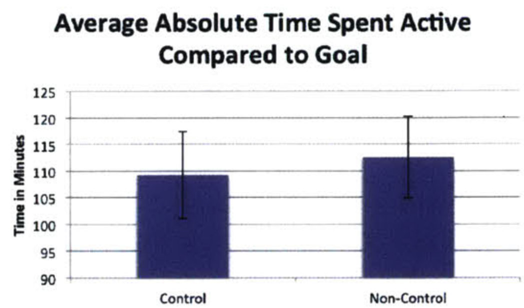




(a) Absolute of average between time spent at home and goal time spent at home

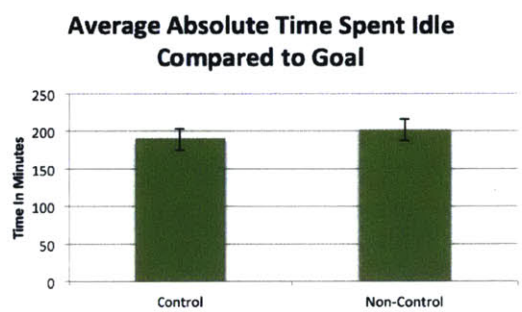
(b) Absolute of average between time spent at work and goal time spent at work





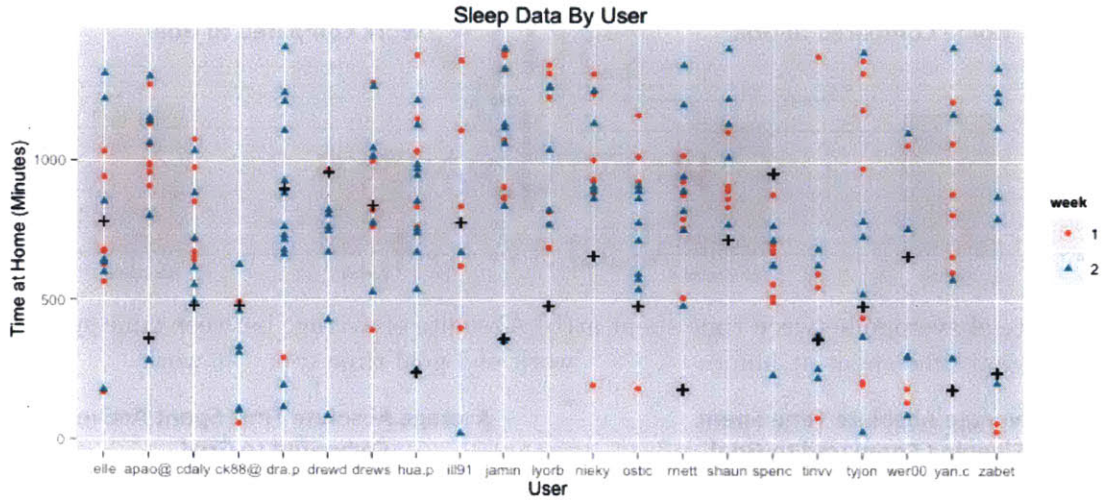
(c) Absolute of average between sleep and sleep goal

(d) Absolute of average between active time and active goal time

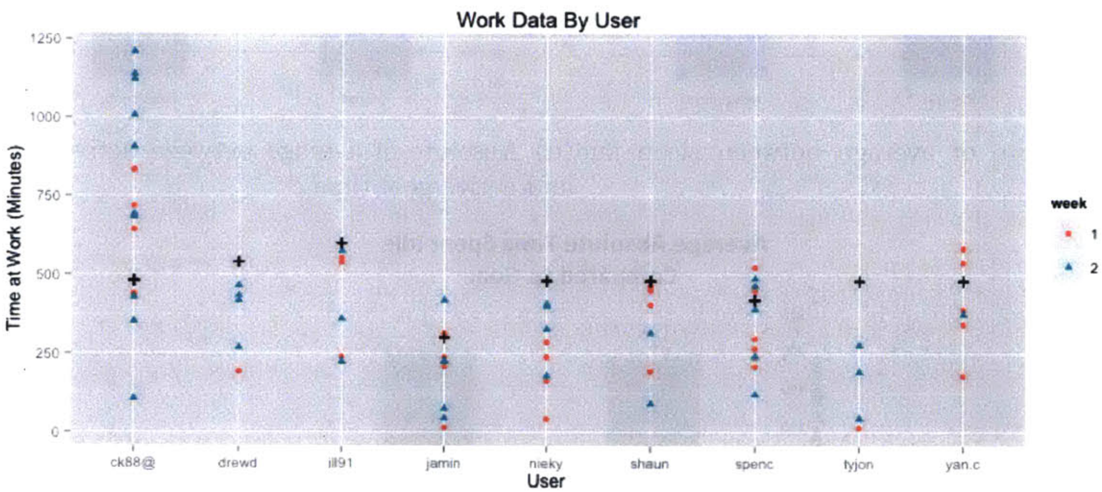


(e) Absolute of average between idle time at home and idle goal time

Figure 3-12: Absolute average of the difference between activity and goal data for the five metrics of Watch System

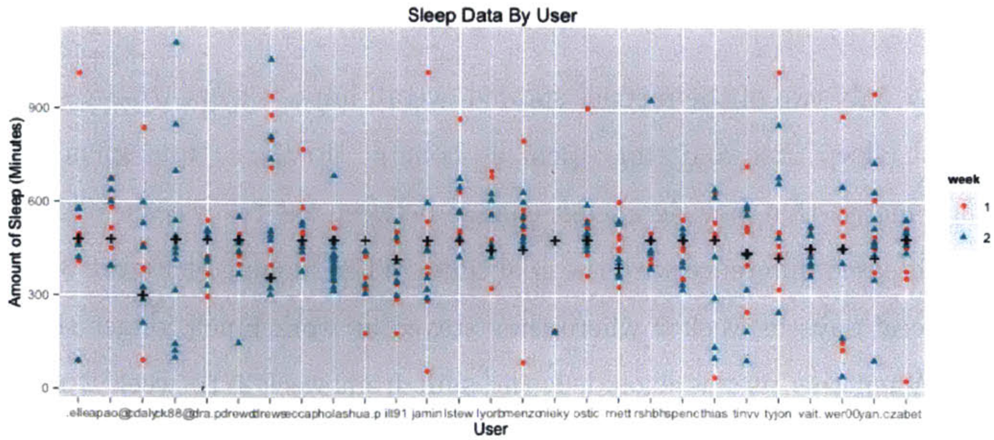


(a) Measured Time Spent at Home

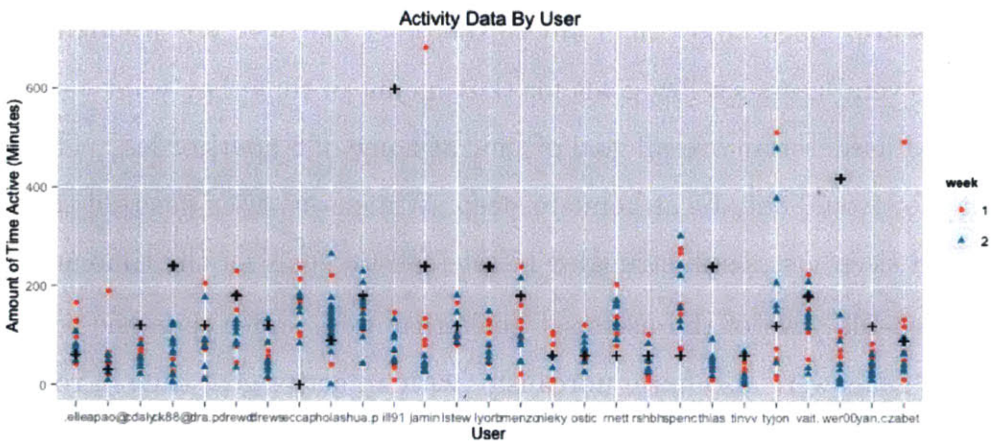


(b) Measured time spent at work

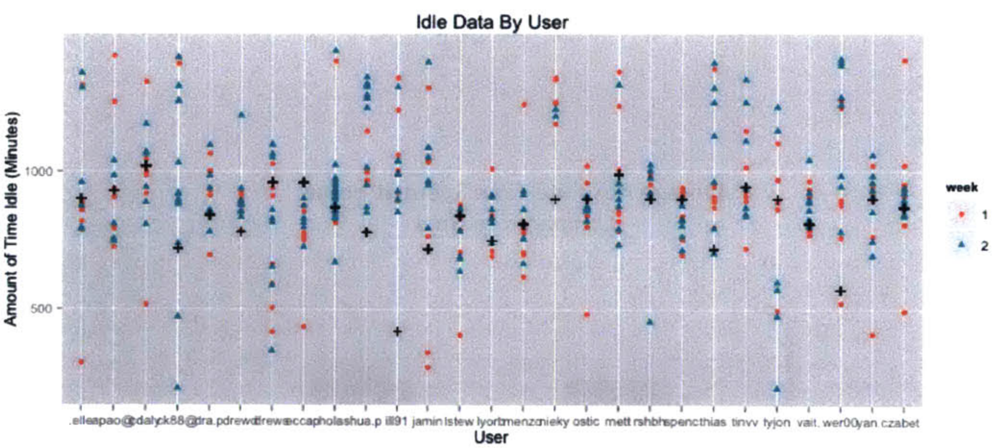
Figure 3-13: Measured data for location metrics by user. "+" signifies a users goal, circles represent first week measured data and triangles represent second week measured data.



(a) Measured time spent sleeping



(b) Measured time spent active



(c) Measured time spent idle

Figure 3-14: Measured data for activity metrics by user. "+" signifies a users goal, circles represent first week measured data and triangles represent second week measured data.

3.3.4 Discussion

In Figure 3-12 only two of the metrics showed overall improvement towards reaching a metric goal (sleep time and time spent at home). However, time spent at work was the only metric that saw significant change between the two weeks. This change showed an increased difference with the goal time. When looking at Figure 3-13a, one user had several days in week 2 where they stayed at work much longer than their goal time. When this user is excluded, the shift from week 1 to week 2 no longer becomes significant.

The absolute of the average difference between week 1 active time and week 2 active time is also much larger than any of the other metrics. When looking at the goal values, several users set low goals for the amount of time they want to spend at home. Several users set an overall goal of time at home of 8 hours or less. This would mean that they would only be at home to sleep. While every user was told that home time included sleep, users still struggled to set realistic goals for the time at home.

In Figure 3-14b, two of the users set goals that were larger than the rest of the group. Their goals were also fairly unrealistic with an active goal time of roughly 10 and 7 hours for these two users as compared to about 4 hours for the next highest user. Removing these two users, as seen in Figure 3-15 brings the average absolute difference to about 80 minutes in weeks one and two, however the difference remains insignificant.

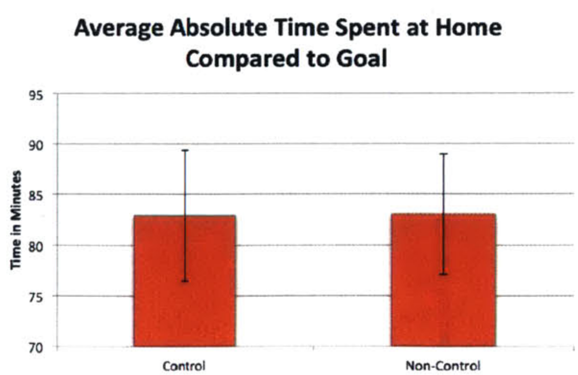


Figure 3-15: Absolute average of the difference between activity and goal data with two users removed

3.3.5 Pilot Study Conclusion

From the data above we can infer that users have a hard time setting realistic time-management goals. This could partly be a result of not knowing how they currently spend their time. The WATCH system was able to passively measure how users spent their time in terms of location and activity. This information was presented to a user as the background of their home screen. This allowed for the information to be presented at opportune moments when a user is not heavily pre-occupied. However without messaging to suggest how to improve, data alone was not effective at motivating users to change their actions so they would be closer to achieving their goals. This study can be revisited with a few improvements. Users should fully understand how to set appropriate goals. Each user should also decide how they want to impact each time goal. This could include spending more or less time at a location or maintain current time habits.

3.4 Conclusion

While data can enlighten users about their current behaviors and even give feedback about how to improve a behavior, without proper messaging the system is likely to be ineffective [53]. When delivered at the right time messages that give specific directions are more likely to result in the desired activity. These messages may be more impactful if they are engaging and include variety. By taking advantage of principles used in persuasive marketing materials, systems could be designed that effectively persuade users. The systems described in this chapter begin the investigation into how messages can be used in personal systems. Future research should investigate ways to aid users in setting realistic yet challenging goals and further investigate challenges in directly applying health behavior messages to wearable technology,

Chapter 4

Experiments in Increasing Motivation

4.1 Introduction

As shown in Bj Fogg's model for behavior change, motivation is a key component for making behavior happen. Education typically increases motivation or reminds the user about the intrinsic or extrinsic rewards associated with performing the desired behavior. However, users of a behavior change system are already intrinsically motivated as they chose to use the system. Therefore, it becomes more interesting to explore methods to introduce or strengthen a motivation through extrinsic forces to investigate whether this can further affect behavior change. This chapter describes a system developed to persuade users to have positive interactions and a second system used to motivate co-workers to take breaks and socialize in their workplace community. Each system is also evaluated with a study.

4.2 Smile Catcher

As previously discussed, our hectic and increasingly digital lives can have a negative effect on our health and well-being. Some authors have argued that we socialize less frequently with other people in person and that people feel increasingly lonely [54].

Loneliness has been shown to significantly affect health and wellbeing in a negative way. Research has shown for years that having a healthy social life is good for well-being and can double life expectancy [55] [56]. Smile Catcher is a multi-user real world game that encourages players to catch as many smiles as possible in a single day or session, in-turn encouraging users to engage in in-person social interactions and get others to smile. Participants wear a Narrative Clip [57] that takes regular pictures of what is in front of them and the system analyzes the pictures captured to detect the number of smiles. The simple implementation of the game allows users to play with only minor changes in their apparel and daily behavior. While other projects have tried to log smiles or even force smiles, not many projects have tried to combine the two in a mobile game for players to enjoy.

4.2.1 Related Work

Previous research has shown that smiles directly correlate to happiness and can even incite happiness in a person [58]. This has resulted in several projects that attempt to measure or force users to smile. All of the projects discussed below primarily focus on inducing smiles in the user. In contrast, the Smile Catcher project encourages the user to induce a smile in another person, which means that they benefit not only from producing a positive affect, but also possibly from emotional contagion [59] and a strengthening of social connections.

4.2.1.1 Mood Meter

The Mood Meter [60] was a MIT project that displayed mood measured in terms of the number of people smiling in a particular location. Camera systems were placed in several areas around MIT's campus. These systems were then used to detect smiles in community members as they frequented these areas. The data collected was used to try to raise awareness of how smiles can positively effect an environment. In contrast with our system, the Mood Meter captured smiles but did not actively encourage users to smile.

4.2.1.2 HappinessCounter

The HappinessCounter [61] was a project that attempted to force smiles through the use of an electronic magnet system attached to a fridge. To open the fridge, the user had to smile. This project was intended for users that lived alone and had limited human interaction. While the HappinessCounter forced smiles and logged smiles, the applications were limited to static locations. Smiles occurring in areas away from the system were not logged. In contrast with our system, the HappinessCounter did not involve any human interactions and dealt with contrived smiles.

4.2.1.3 Emotional Flowers

Emotional Flowers [61] is a game that also utilizes the player's emotion. A user can play with several people in their social circle. A webcam on each player's computer monitors his or her facial expression. Facial expressions classified as happy or surprised allow their personal flower to grow and change its color. Negative emotions result in their flowers shrinking. Players attempt to competitively grow the most flowers. Emotional Flowers also utilizes gamification to encourage happiness, however it does not encourage positive interactions among players in the game. Besides competing against each other, there was no advantage to interacting socially. The short interval at which a smile would be captured resulted in players forcing smiles several times an hour. This resulted in players quickly becoming tired of having to smile at such frequent time periods.

4.2.2 Game Implementation

Most users play Smile Catcher with the Narrative Clip, however the game can be modified for other wearable technologies that include a front facing camera such as Google Glass. Game implementation consists of three parts: image capture, image processing and feedback. A session is typically played over the course of the day; however the game can last indefinitely.

On the Narrative Clip, a game or session does not need to be started as images are taken automatically every 30 seconds (this was a technical limitation of the Narrative Clip at the time of the study). Proper capture depends on correct placement of the camera and the moment being captured by the camera. To process the images captured by the Narrative Clip, players use an online application to upload images. Each image is processed using the Face++ API [62]. The images were analyzed for faces and subsequently smiles. If a smile is present, the image is saved for the player to review. Other images are not saved as the player already has the image saved on their personal hard drive and on the Narrative Clip server. The web application saves their daily counter of smiles and shows trends.

At the end of the day, players are given a score of number of smiles caught calculated for the day or agame counter (if multiple days). This score can be sent to others to incite a long-term competition. Hopefully, this score will become as important as how many steps were taken or how many calories were consumed during a day.

4.2.3 System Design

The Smile Catcher system utilizes the Narrative Clip as a trigger to engage in a positive social interaction. Users are reminded of the goals whenever they see the clip as well as when others ask what the device is used for. This trigger results in a positive interaction. The user is immediately rewarded with the positive affect that comes when spreading joy. They are later rewarded again when they are able to review these positive moments. The game was evaluated with a two-week study that sought to determine if this system can increase positive social interactions and thereby people's well-being.

4.2.4 Study Design

4.2.4.1 Participants and Methods

Our game was evaluated using a two-week study with thirty-three participants. Participants' ages ranged from 19 to 67 (14 males and 19 females). The participants were recruited from a database of willing study participants. Each were told they would be required to wear the Narrative Clip for two weeks and would upload the images at the end of each day. So that their first week could be used as a baseline, they were not told the intention of the study. The only criterion for participation was daily access to a computer with reliable Internet.

A participant was considered to have successfully completed the study if they fully participated throughout the two weeks. Twenty participants successfully completed the study. Full participation consisted of uploading all of the images captured at the end of each day. Successful participants uploaded at least 100 images each day; ten participants did not upload a sufficient number of images (less than 100 over the two week period). Three participants lost their tracker and therefore were unable to participate for the full two weeks.

4.2.4.2 Pilot Study Procedures

Prior to starting the pilot study, each participant attended a thirty minute orientation session where they were given the Narrative Clip, shown how it worked and given an overview of the website they would be using throughout the duration of the study. Each participant was given suggestions for how to avoid privacy issues. They were permitted to remove the device if an individual felt uncomfortable or they were in a private environment.

Each night, participants visited the website to upload their images. The first week of the pilot study the website was called Life Logger and did not mention anything about smiles. Participants were asked several questions in regards to their mood (see Figure 4-1). Once they completed the quick survey, they would gain access to a form to upload their images. During the first week, this was the final step needed to successfully complete the daily requirements.

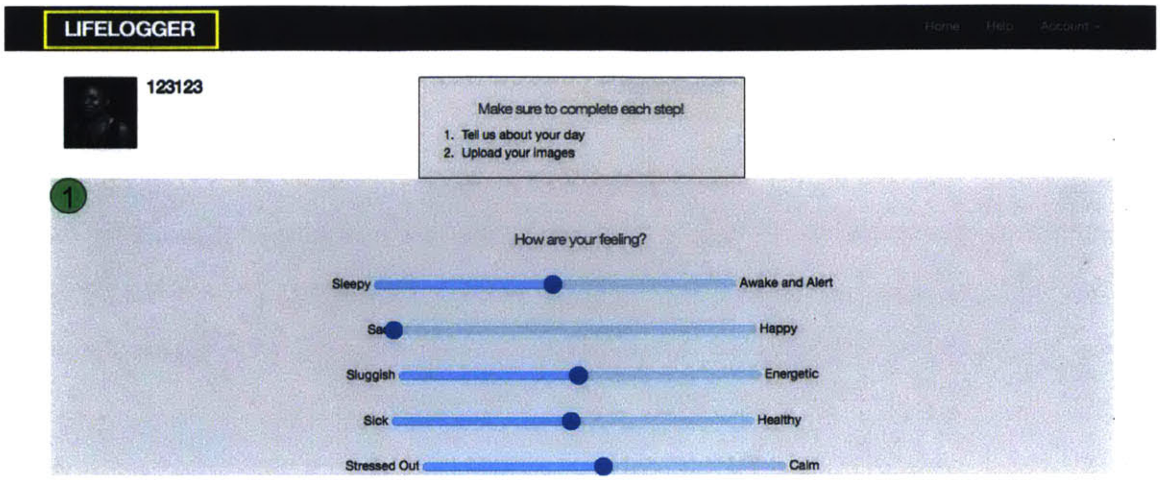


Figure 4-1: Baseline week upload page and questionnaire

At the start of the second week (Figure 4-2) of the pilot study the website name changed to be called Smile Catcher. A smile goal was visible at the top of the website as well as a tracker that reflected any past smiles caught. This goal was the same for all participants and set by averaging daily smile totals from test participants. Participants once again had to report on their day before uploading their images. After uploading, the users were able to review five randomly selected smiles caught in previous days and were once more asked about their mood. After completing the survey, they were shown graphs illustrating trends in the number of smiles they caught. Participants were not rewarded for reaching their smile goals. These goals were set for each user.

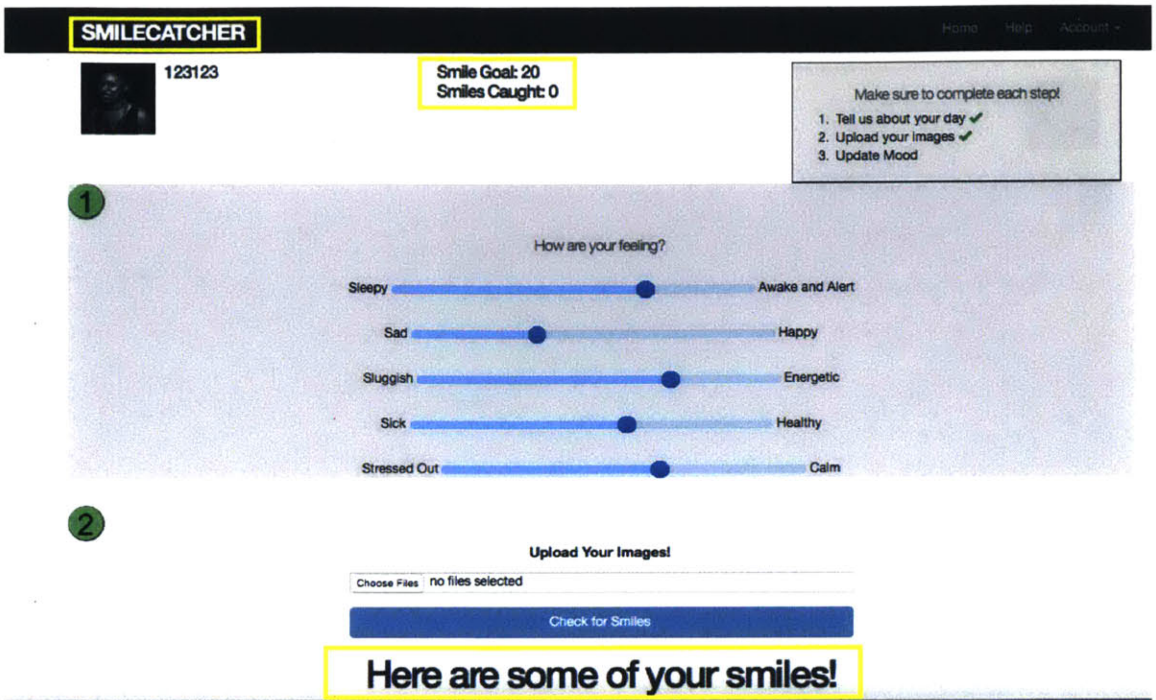


Figure 4-2: Week 2 upload page

Figure 4-3: Some of the smiles caught by participants

At the conclusion of the pilot study each participant attended a thirty-minute closing session. During this time they returned the Narrative Clips, completed a post survey and were asked about their experience playing the game.

4.2.5 Results

We analyzed the total count of smiles caught to determine if participants had more positive social interactions during the pilot study. The first week of the pilot study was treated as baseline. When introducing a new device to users, a novelty effect may occur [63]. Our participants were not initially aware of the purpose of the pilot study. If a novelty effect were present, any peak over the baseline week would further confirm that we were able to influence social behavior.

Evaluating the smiles caught by the twenty successful participants, we found, with significance ($p<.05$), that participants caught more smiles during the second week of the pilot study. Users caught an average of 6.25 smiles during the baseline week and 10.85 smiles during the following week. Those that caught more smiles in the second week increased the number of smiles caught by 9 images on average, while other participants caught an average of 3.57 fewer smiles in the second week.

During the closing session most participants were able to report on smiles that were caught or missed. Participants developed tactics to increase their chances of catching smiles. Several participants noted making efforts to engage with strangers in an attempt to gain more smiles. Other participants chose to smile more at individuals hoping they would reciprocate the gesture.

While users were able to mentally recall smiles in their day, this did not have any significant effect on their self-reported mood. This may be due to participants only submitting their mood at the end of the day. Previous work has shown that recall of mood can be inaccurate [64]. Our participants may have more accurately submitted their current mood and not an accurate measure of the day's mood. While there were no apparent trends between number of smiles and the mood data collected at the end of the day, 13 participants commented on how the number of smiles they caught affected them either positively or negatively.

Not catching smiles was disappointing to many participants; this may be due to limitations of the system. Due to the nature of the Clip, it was easy for smiles to be missed. Participants were very aware when a smile was not caught by the camera and expressed frustration. Misses were either caused by an error in the detection software or simply not capturing the moment due to the capture rate of the Narrative Clip. This demonstrated that users were independently searching their images for smiles and that they were also able to mentally log smiles during the day.

Users found other trends in the images they captured during the day. Several users commented on the majority of their images being of their computer screens. They were shocked by the amount of time they spent in front of a computer and the lack of time spent having human interactions. They also expressed disappointment in the lack of smiles they caught during the pilot study. Three participants so much enjoyed the process of capturing their day and the introspection it provided that they decided to purchase their own devices so they could continue after the pilot study ended.

4.2.6 Discussion

4.2.6.1 User Feedback

Through user feedback we were able to identify several areas where the game could be improved. It was clear that for long-term adoption, accuracy of the game and ease of use would need to be improved. Accuracy can be improved in further iterations by utilizing a more accurate API to analyze images. Newer versions of the narrative clip have the ability to change the image capture frequency. Increasing the capture rate decreases the chance that a smile would be missed. Outside of the pilot study we would like participants to be able to play without having to upload their images. However, there is currently no open API to develop directly with the Narrative Clip. Women also expressed frustration in trying to find an effective location to wear the clip. Most examples given by the company highlight men and clothing typically worn by men on a regular basis. Women tend to not have a flat surface on their chest where the clip could be easily pinned.

4.2.6.2 System Limitations

Using a computer vision algorithm to determine smiles results in errors in detection and false positives. Several participants used public transportation daily. Advertising used on public transportation often uses images of happy people. If this were in the participant?s field of view, it would result in several "catches". However these catches accounted for less than 10% of all smiles caught and only three participants caught smiles of this nature. We allowed these to be counted toward the overall total, as emotional contagion is effective in print materials [65].

4.2.7 Pilot Study Conclusion

Through a two-week pilot study we were able to show that gamification could be used to persuade users to attempt to have more positive social interactions in their day-to-day lives. By simply tracking the smiles users witnessed during the day and giving a goal, users significantly changed their behavior. We expected users to attempt to target people in their social circles when attempting to catch smiles. However, many users reported engaging with people outside their social circles. We also expected users to be aware of the positive moments in their day. This proved to be the case, but users were also made aware of their lack of social interactions. By simply asking users to catch smiles, they became more cognizant of the smiles in their day-to-day actions. The ability to positively recall memories has been linked to an increase in overall mood and gratitude [66] and therefore an improvement in well-being.

Overall, we were able to positively persuade pilot study participants of all ages to change how they will interact with individuals inside and outside of their social networks through the use of a new trigger. The structural aspects of a behavior trigger utilized by the Smile Catcher is summarized in Table 4.1. While game design was used to encourage behavior change, participants were not rewarded for completing their goals, this type of self-motivation will lead to long-term change.

	Motivation	Attention	Timing	Message Content
Week 1	X	√	√	X
Week 2	√	√	√	√

Table 4.1: Summary of trigger usage for Smile Catcher system

4.3 The Challenge

The influence that additional motivation can have on behavior change was explored further in the development of another system called The Challenge system. Individuals who work in sedentary occupations are at increased risk of a number of serious health consequences. Physical activity has been repeatedly shown to be important for maintaining overall health and preventing mortality [67]. For individuals who maintain a sedentary lifestyle, the risk of death from coronary heart disease may be as much as 1.9 times that of more active individuals [68]. A meta-analysis of 18 studies found that those with the greatest amount of sedentary time had a 116% increased relative risk of diabetes, and a 147% increase in the relative risk of cardiovascular events [69]

This is especially problematic for the increasing number of people who perform sedentary occupations. Large scale studies (e.g. [70]) have revealed that people who perform more vigorous physical activity at work have a significantly decreased risk of mortality compared to those in sedentary occupations. For these reasons, researchers studying physical activity at work recommend that office workers take breaks from sitting at their desk to engage in physical activity, in order to interrupt prolonged sedentary periods [71].

Individuals in academic settings such as a graduate research lab are vulnerable to these risks, given the long hours they spend working in a sedentary position. The workload of many graduate students is such that they may be unwilling to take time away from their desk to participate in physical activity. In addition to necessitating extended periods of sedentary work, the heavy workload may cause students in such labs to become socially isolated, because they do not feel they can spare time away from research to participate in social activities. Unfortunately this tendency can be extremely harmful, as social isolation increases the risk of a range of mental health issues, including depression, anxiety, anger, diminished optimism and self-esteem, impaired cognitive performance, and suicide [72].

With Natasha Jacques of Affective Computing in the MIT Media Lab, we designed a tool and experiment aimed at decreasing sedentary activity and promoting social connections in a workplace environment. Our system asks participants to sign up for short physical challenges and pairs them with a partner to perform the activity. Pilot testing of the system was performed at the MIT Media Lab, where participants' overall activity levels were monitored with a pedometer-based activity tracker to assess the effectiveness of the interventions.

4.3.1 Related Work

There is a broad body of research on interventions designed to improve physical activity, which we will not attempt to cover here; several meta-analyses are available in [73], [74], [75], and [76]. Despite the interest in increasing physical activity, it remains a difficult problem; in many studies, nearly 50% of participants drop out or fail to adhere to the intervention [74]. Individuals' adherence tends to drop when they view the exercise less positively, or as too physically demanding [74].

Still, interventions to increase physical activity and reduce sedentary behavior can be effective, especially when targeted directly at the workplace. A meta-analysis of workplace-based interventions showed that they significantly increase physical activity and reduce the risk of diabetes, and are actually more effective when delivered at the workplace rather than elsewhere [73]. Further, they have the additional benefits of significantly improving work attendance and job satisfaction, as well as reducing job stress. The authors hypothesize that the increased effectiveness of workplace-based interventions may be due to the ability to set behavioral norms, as well as increased group support for the program [73]. Indeed, group cohesiveness has been shown to improve adherence to physical activity interventions [77].

Pedometer-based activity interventions are also highly effective. A meta-analysis of 26 studies demonstrated that when wearing a pedometer, reliability increases physical activity, on average by 2491 steps per day [78]. The increased activity due to wearing a pedometer led users to decrease their body mass index (BMI) by 0.38 and systolic blood pressure by 3.8 mm, on average. A workplace-based intervention using pedometers and a mentorship program, which taught participants about goal setting and the benefits of physical activity, showed an increase of 3451 steps on average [79]. Participants in this study also reduced their BMI, waist girth, and resting heart rate.

Work has also been done on reducing social isolation, however the effectiveness of such interventions has been questioned. In a review of 20 studies, Findlay argues that there is little evidence that interventions designed to reduce social isolation among elderly adults actually work [80]. Of the interventions that were successful, most attempted to use existing community resources to build community capacity [80]. Another systematic review found that successful interventions tended to be group activities, while unsuccessful interventions tended to provide one-on-one social support or advice [81]. Therefore in designing interventions to reduce social isolation, it appears to be more effective to target communities and groups, rather than individuals.

4.3.2 System and Intervention Design

We have developed a behavioral intervention technology to specifically target the needs of the Media Lab community. We have chosen to focus on easy physical activities that can be done for short periods of time within the lab itself, since we believe participants would be most willing and able to participate in these. Given the importance of social influence in behavior change [82], we use social obligation to promote participation and reduce social isolation.

Our system consists of an online tool that allows individuals with MIT Media Lab credentials to register and sign up to participate in up to three competitions for table tennis, foosball, and walking. Once a participant has joined a competition, she is eligible to receive *challenges* in the form of an email asking her to engage in the activity. She can accept the challenge, reject it, or simply let it expire (it will expire in five minutes). If the challenge expires or is rejected, the system will randomly select another participant who is enrolled in the same competition and issue a challenge to them, until someone accepts or all possible options are exhausted. We chose to have the system randomly select participant pairs to promote social connectivity and prevent users from simply playing with those they already know (initial conversations with participants revealed that most tend to have a small number of people with whom they regularly engage in these activities). Once two participants have accepted the challenge, the system will email both asking them to meet in a specific location, and include the name, photo, and email address of the other participant in the message.

We take a gamification approach to encourage participation, by awarding points when participants successfully complete challenges. These points translate to their score in the competition for each activity; i.e. if a participant completes a foosball challenge, she will be awarded a point in the foosball competition. There is also an overall competition, in which participants are awarded 10 points for every challenge they complete (of any type), and 1 point for every 1000 steps. At the end of the competition, the participants with the most points in each of the competitions (table tennis, foosball, walking, and overall), earn a prize of their choice. Prizes are either an activity tracker such as the FitBit Flex or equivalent monetary compensation. Because competition can be intimidating for some users, we award prizes simply for playing games rather than winning them. We also include going for a walk as a possible challenge activity, since it requires no previous experience to perform.

The most effective behavior change mechanisms make the behavior simple and easy to do, provide motivation, and trigger the behavior at the correct time [83]. Our approach addresses all three of these criteria. By targeting short activities that can be done indoors in the lab, we make it easy for participants to take a quick break and perform them, without losing much time away from work or having to navigate possible adverse weather conditions in order to exercise. Activities like table tennis and foosball cannot be performed without a partner; therefore by helping people find a companion to play with, the system makes these behaviors easier as well. The social aspect of being paired with someone may also increase motivation, as social obligation can be a powerful factor in adhering to an exercise routine [84]. Further, social influence is also an effective persuasion technique in behavior change [85]. When an individual sees that many people around her are participating in an activity, she will be more inclined to participate as well [85]. This experiment was able to leverage this principle, both because users of the system are able to see and meet with the other users of the system, but also because the table tennis and foosball tables are set up in a highly visible and trafficked part of the Media Lab.

Finally, we implemented two mechanisms for triggering participation in the targeted activities. First, users can log on and issue their own challenges; which causes the system to find another user who is willing to play a game with them at that moment. In this way the person requesting a partner is actually providing the trigger for other participants in the experiment. The experimenters can also manually trigger matches. The mechanism in this case takes a similar form; the system will cycle through users interested in the given activity until two are found who are willing to play.

Several design choices were made that were intended to increase the usability and robustness of the system. A previous system at the lab designed to pair people to play table tennis at scheduled times fell out of use because participants would fail to attend matches, leaving their partner waiting at the table with no explanation. This led to a highly negative user experience. Our tool improves on this with several mechanisms. If a participant has accepted a challenge but the system has not been able to find a partner for them within 20 minutes, it will send another email asking them to confirm that they are still interested in playing. That way, whenever a challenge is arranged, both participants have confirmed their availability within the last 20 minutes. When a challenge is arranged, we include the name, photo, and email of the challenge partner in order to increase accountability and prevent no-shows. However, it may still be possible for a participant to fail to show up for a challenge in spite of the fact that the partner knows her name; for this reason, we also include a mechanism for reporting that a partner did not attend, which deducts a point from the partner's score. Participants themselves can also report that they are not able to attend an arranged match, in which case the system will email their partner and inform them that the match has been canceled.

Based on expert advice, we also added a feature that sends the participants a weekly update email informing them of the current leaders in each competition and their points. The email then goes on to describe how many points the participant herself will need to reach third place in each competition. Because in some cases only a few challenges need to be completed to reach this point, it was thought that these emails would encourage participation from users who had not previously engaged with the system.

4.3.3 User Study

To assess whether our tool had a beneficial effect on promoting activity and social companionship, we conducted a 5-week-long experiment in which participants were monitored via Jawbone Up activity trackers. The trackers can monitor not only each participant's daily step count, but also the maximum amount of time spent idle each day. This data allows us to assess the hypothesis that our system reduces sedentary activity and promotes physical activity.

Participants were recruited through an email to a Media Lab mailing list, with the criteria for inclusion being that the participant had valid Media Lab credentials. Enrollment for the full study was limited to 48 participants based on the number of available Up trackers, but additional users were allowed to register for the site and sign up for challenges; a total of 59 users registered. Out of the initial 48 participants who received activity trackers, 8 lost their trackers, one tracker malfunctioned, and 5 other participants kept the tracker but did not actually use it to log activity information. Therefore we were left with a total of 34 participants for analysis.

In order to assess whether our interventions had an effect, we used a within-subjects study design in which participants served as their own control. Prior to the first week of the study, participants signed an informed consent form, received their Jawbone Up tracker, and completed a short survey asking them about the number of people from the lab they interact with on a weekly basis, their current level of physical and sedentary activity, and whether they already owned a fitness tracker. Once the study commenced, the first week acted as a baseline week, in which participants wore the tracker but use of the challenge website was limited; no challenges took place and no points were awarded in the competition. Starting in the second week, these features were enabled; challenges were issued by both the experimenters and participants (approximately 5% of challenges were issued by the experimenters), points were awarded, and participants received weekly emails showing the competition leaders and the points required to win. After three weeks, the competition ended, although the study did not. Participants were asked to continue wearing the trackers for another week, to determine if their activity level remained high after the interventions ended. At the end of the study, participants completed the same short survey about activity levels and social diversity, and an additional follow-up interview about their experience with the system. Participants were compensated up to $90 for full participation in the 5 weeks of the study.

We hypothesized that our intervention would have three effects:

1. Decrease sedentary time below baseline
2. Increase activity level above baseline
3. Allow participants to meet new people in the lab

Note that simply receiving the activity tracker could act as a confounder in our study, since people tend to perform more activity when they initially begin to use an activity tracker because of the novelty effect [78]. However, this confounder would increase the activity in our baseline week and make it harder for us to detect an effect of the intervention, so any effect that remains is likely more robust.

4.3.4 Results

In this section we will first present quantitative results about engagement with the system in general, then address whether the data support each of our three hypotheses, and finally present a detailed analysis of the qualitative feedback we received on the design of the system.

Figure 4-4 shows the number of challenges issued for every day of the study. The first week is the baseline week, during which the challenge feature was disabled and no challenges could take place. The competition began on March 26th and continued over the next three weeks, ending on April 15th. Engagement with the system continued

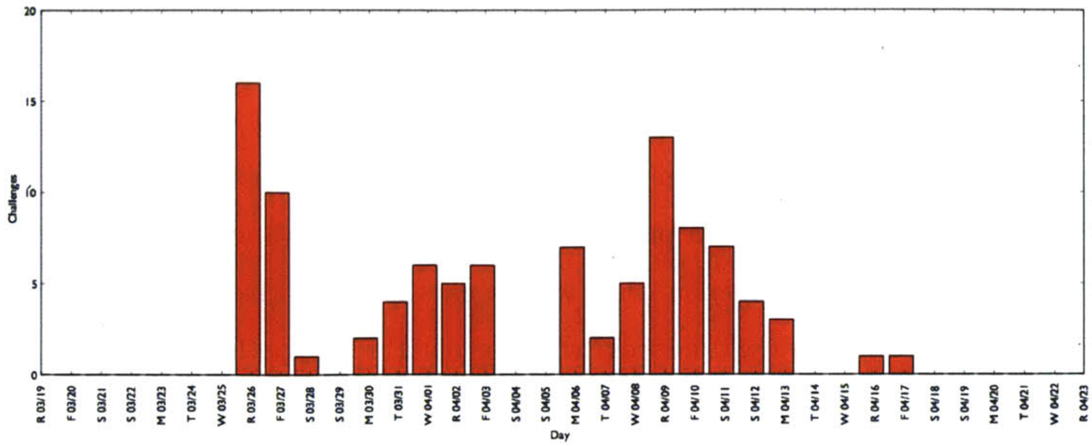


Figure 4-4: The number of challenges completed for each day of the study.

to be monitored during the following week after the competitions ended. We note that some challenges still took place, which is encouraging. In total, 101 challenges were completed over the course of the study. On average, participants received 11.12 challenges ($SD = 8.66$) and completed 2.53 ($SD = 4.05$).

From this data we note some interesting trends. Normally, few or no challenges took place on Saturday or Sunday, presumably because participants were at home rather than at work. This trend is reversed on April 11th and 12th. We believe this may be due to a high-pressure event that significantly impacts the lab community as a whole. During this event (called *Members' week*), students present their work to interested companies in industry. Many students work long hours in preparation for Members' week in order to have demos of their work ready for presentation. The demos took place on the 14th and 15th of April, during which time the table tennis and foosball tables were inaccessible due to the event (and as Figure 4-4 shows, no challenges took place on these days). We believe the increased challenges on the weekend of April 11th reflect the fact that many students were uncharacteristically at the lab working during this time, in order to prepare for Members' week. We will see later that preparation for Members' week may have affected students' activity in other ways.

The level of engagement with the system was highly variable between participants. While every user received at least one challenge, only 18 ever actually completed a challenge. The median number of challenges completed was 1.0, but there were super users who participated in a much larger number of challenges; in one case, as many as 17. Figure 4-5 shows the distribution of challenges completed by user. These data suggest that while some participants seem to particularly enjoy using the system, many were reluctant to participate in challenges. We will explore possible explanations for this phenomenon when we discuss the qualitative feedback provided by participants. First, we will address whether our three hypotheses are supported by the data collected.

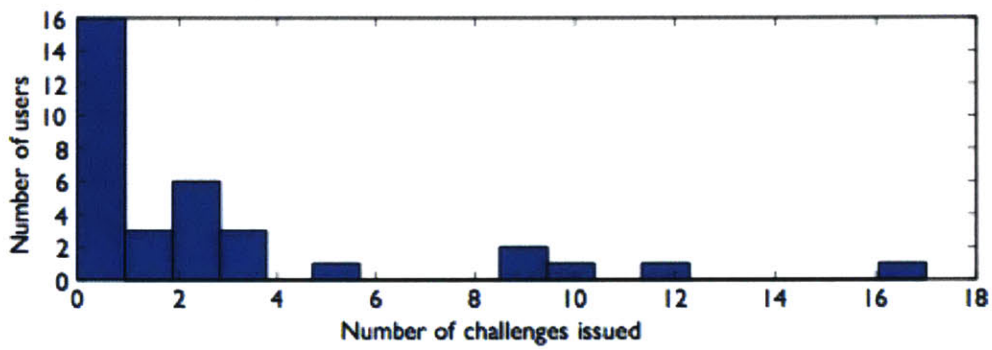


Figure 4-5: Histogram of the number of challenges completed by each user

4.3.4.1 Hypothesis 1: Sedentary time will decrease

One of the primary goals of the Challenge system was to try to prevent prolonged periods of sedentary activity that occur when students work at their desk without taking a break. We hypothesized that the short and easily accessible nature of the challenges would persuade students to leave their desk instead of remaining in a sedentary posture. The data appear to support our hypothesis. There was a significant negative correlation between the number of challenges completed per day and students' average maximum amount of time spent in continuous sedentary activity, $r(36) = -.351$ (pearson R score), $p < .05$. In other words, the more challenges that took place in the lab on a given day, the shorter the time those students spent in prolonged sedentary activity. This result is encouraging, and suggests our intervention may be appropriately targeting the behavior we want to change.

Yet engagement with the system remains a problem. If we examine whether participants' idle time decreased from their baseline week and include all 34 participants (even the ones who never completed a challenge), the results are not significant. This result is expected. If completing more challenges is associated with less sedentary activity (for which we have evidence as detailed above), then participants who do not complete challenges will not benefit from this intervention, and will not have a decrease in their sedentary activity.

To test this theory we restrict our analysis to only those participants that completed at least one challenge. We applied a Tukey's HSD (honest significant difference) test and found that for these participants, total idle time[1] decreased significantly ($p < .05$) below their baseline in each week of the intervention and after the study ended, except for the second week of the competition (Figure 4-6 illustrates this result). The lack of result for week 3 is interesting, and is a trend that occurs throughout our analysis. Note that week 3 of our study overlaps heavily with the week before Members' week; during this time, students are often working long hours in preparation for Members' week, which could provide an explanation for their increased sedentary time. A quote from one participant confirms, "the week before Members' week is more stressful than Members' week" (P33). Despite the lack of result for this week, overall these data suggest a general trend of decreasing sedentary activity for participants who used the challenge system, although the same trend does not exist for the whole group of participants.

Participants also self-reported spending fewer total hours in uninterrupted sedentary activity by the end of the study. In the pre-study survey that took place before March 19th and in the post-study survey after April 24th, participants were asked, "On average, how many hours do you typically spend working at your desk without taking a break?". Figure 4-7 shows the difference in the value reported on this question for each participant. We find that most participants report decreased sedentary activity, leading to a significant difference over all participants, $t(34) = 2.54$,

[1] In order to compute measures of idle time, we excluded days on which participants did not track their sleep time using the Up tracker, because we were concerned that in this case sleeping could be confused with idle time.

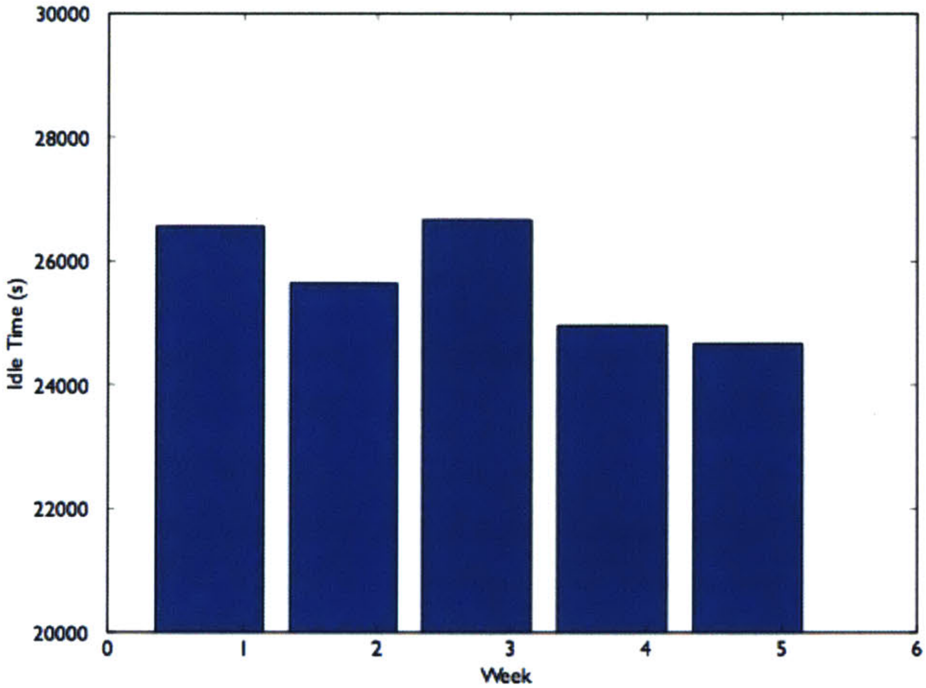


Figure 4-6: The total idle time per day averaged over each week of the study for participants who completed at least one challenge

$p < .05$. Interestingly, participants seem to be more optimistic in reporting their sedentary time than the results from their Jawbone Up would suggest. This may be a case of observer-expectancy effect [86], but may therefore suggest that participants' awareness of the importance of reducing sedentary activity has increased.

4.3.4.2 Hypothesis 2: Activity will increase

In addition to reducing sedentary activity, we have found that completing challenges is related to more total active time per day. There was a significant positive correlation between the number of challenges completed on a given day and the average total active time of participants, $r(36) = .331$, $p < .05$.

However, as we saw in the last section, even if completing challenges is related to increasing activity, this level of analysis cannot tell us if all participants are benefiting from the intervention. Therefore we examine the step count results for all 34 participants to see if steps increased over the initial baseline week. Using a Tukey's HSD

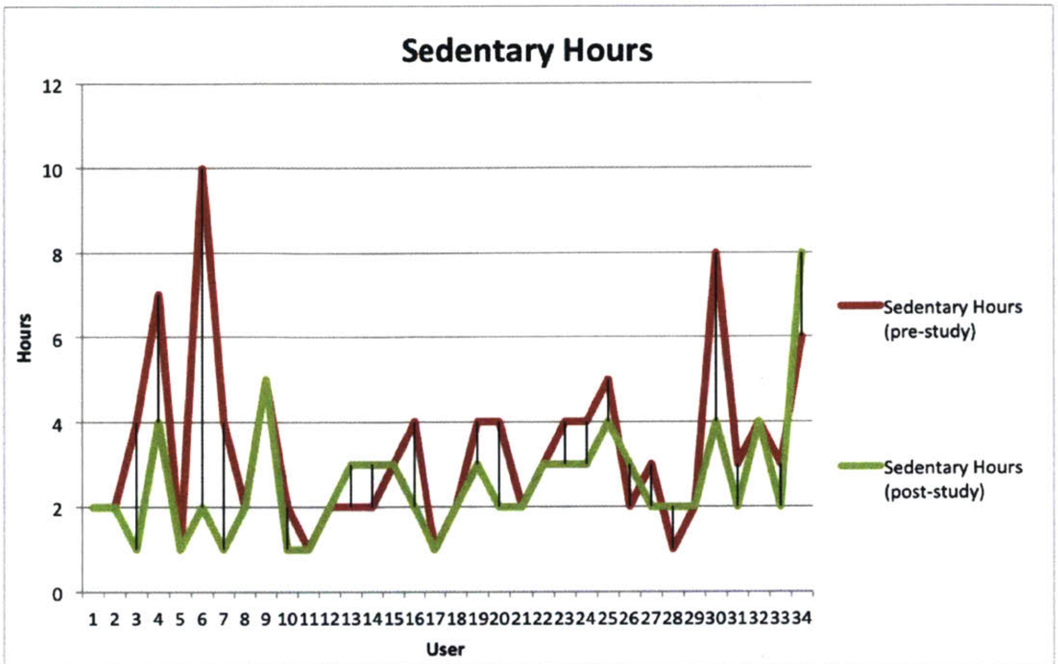


Figure 4-7: Change in self-reported sedentary activity for all 34 participants. Pre-study values are in red, post-study values are in green.

test, we find that steps increased significantly ($p < .05$) over the baseline during the first and last week of the competition (Figure 4-8 graphs this result). A similar result was found for total active time; it increased significantly in the first and third week of the competition, although not in the second. This pattern of week 3 of the study not showing significant results mirrors the trend found above for sedentary time. We again hypothesize it may be due to students being overworked before Members' week.

Interestingly, trends of increased activity and step count are evident when computed over all users, even though not all users completed challenges. We believe this is due to the motivating effect of the competition itself. Even users who did not complete challenges were still eligible to compete in the overall competition based on their step count, and received weekly update emails informing them of how close they were to winning. In fact, none of the three participants who won the overall competition completed challenges, but won purely based on step count[2]. Further, if we compare users who participated in zero challenges to more engaged users (those who completed at least three challenges over the course of three weeks), we see that the group with

[2]This may suggest we may need to alter the point structure in future iterations of this study

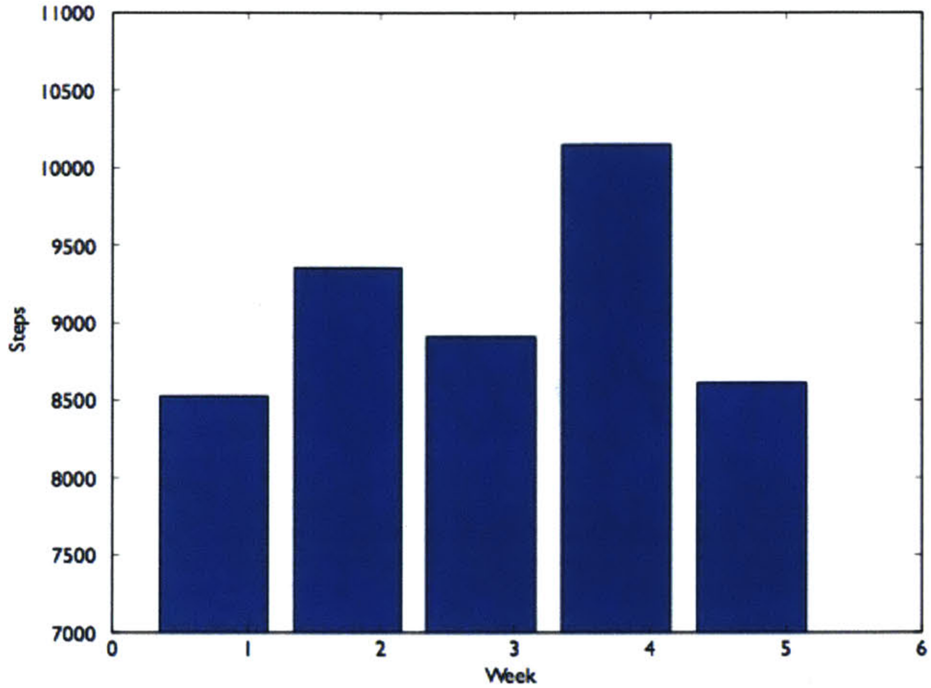


Figure 4-8: Steps per day averaged over each week of the study for all participants

zero challenges had significantly higher step count per day, $t(284) = 8.20$, $p < .01$, and significantly higher total active time per day, $t(284) = 6.46$, $p < .01$. This suggests that not only were participants who performed zero challenges increasing their steps and activity above their own personal baseline, but they were also increasing their activity above participants who gained points in the competition through challenges. The most likely explanation for this phenomenon is that participants who did not wish to do challenges were still highly motivated by the competitive aspect of the study.

4.3.4.3 Hypothesis 3: Users will meet new people

One of the main goals of the Challenge system was to allow users to meet new people within the lab, in order to broaden their social groups and reduce social isolation. In the end-of-study interviews, we found that participants did report meeting new people as a result of using the system, and that some appreciated this feature. For example, P26 said, "I met new people I didn't know before and I liked that", and P49 stated, "I think it's always a nice opportunity to meet somebody". P12 reported, "It worked just how it was supposed to. I felt like I'd been sitting at my desk too much, issued a challenge, met a new person, and spent a thoroughly enjoyable 20 minutes". Even in cases where the system paired participants with someone they had already met, it may still have helped to reduce social isolation. P41 explained that he was paired with one of his lab mates that he did not socialize with frequently before the study. He stated, "We will probably just ask each other to go walking now. Which is good because we weren't doing that before". Another participant who was new to the lab community said that as a result of the study, not only did he meet new people, but he also "learned other things about the Media Lab, about the building" (P49). These experiences demonstrate that the system can help lab members to make new social connections or strengthen existing ones.

However, in conducting interviews we learned that in many cases participants disliked the idea of meeting someone new. They stated that they would have preferred to be able to challenge "someone you know rather than a stranger" (P27), and that they find it is "more fun going on walks and playing ping-pong with close friends" (P17). Some participants had a previously established group of people to play with, and were not interested in meeting anyone else for this purpose; e.g. "I have specific people I play with" (P24). The random-assignment aspect of the challenges may actually have deterred some participants from participating in them. As P13 exclaimed, "It's too scary to play pingpong with a stranger!"

4.3.4.4 Qualitative feedback on improving system design

The end-of-study interviews provided us with rich, qualitative data about participants' experience with the system and ways in which we could improve the design to facilitate a better user experience. Users offered many insightful suggestions; for example, having the system arrange doubles games because, "It gets more intense when you play with more people, and it's easier to get to know someone" (P33). Having other milestones worth points in the competition was recommended as well.

P32 suggested, "You win like a little badge if you walk to a certain location". This is an idea we would like to incorporate, especially after hearing from P30, "I didn't issue a walk challenge because I wanted to walk but I was concerned it might take too much time". Another valuable idea concerned motivating the group of participants as a whole, as well as the rest of the lab. P49 proposed placing an interactive screen in the atrium (a well-trafficked public place in the lab which contains the table tennis and foosball tables), in order to display active challenges and aggregated activity results.

Participants had a lot to say about the challenge mechanism itself. There were several comments about the fact that it often took too long to find someone to play with after a challenge was issued; e.g. "It was taking longer and longer to get people" (P46), and "I had to leave before it found someone" (P38). This experience could be discouraging for participants. P33 felt, "it's just sad when it's like, '20 minutes have passed, are you still available?'". Improving the speed of the current design would require shortening the 5-minute expiry period (recall that the system cycles through users one at a time, giving them 5 minutes to accept the challenge before moving on to the next one). However, participants also complained that the 5-minute expiry period was too short. There were several comments such as, "When I tried to accept a challenge the time had expired" (P53) and "It was already expired by the time I tried it". Despite the fact that participants could accept an expired challenge (they just would not be given priority if the person who was currently being challenged also accepted), this apparently was not clear enough. P55 astutely pointed out, "Your email only updates every 15 minutes", so expecting participants to respond within 5 minutes to an email based challenge is unrealistic. P21 confirmed, "I missed answering challenges several times".

This feedback helped us determine the best ways to overcome the shortcomings of the system, for example by having multiple methods for communicating with users. Several participants emphasized this point: "I would prefer [the challenges] to be sent to my phone" (P14), "if it came to SMS I would be more available" (P13), "Text messages are on the top of my priority list" (P15), and "if changed to text messages I would possibly keep using [the system]" (P19). Given this level of demand for sending challenges via text message, adding this functionality is our top priority for improving the system.

We are also determined ways to revise the challenge mechanism to reduce waiting without reducing the expiry period. Rather than cycling through participants one at a time, we can send a blanket email to all participants who might be interested in playing. We originally rejected this idea because we did not want to 'spam' participants with too many emails. Yet some participants suggested a mechanism for "seeing if people are available right now to play ping pong" (P38). This could consist of a type of "online mode" (P32), in which you can "indicate you're available right now for challenges" (P32). Participants could set the system to 'online mode', and when this is activated they would receive all challenges immediately. This could make it easier for participants issuing challenges to find a partner, but still prevent busy participants from being spammed.

Another method for dealing with latency issues is scheduling challenges ahead of time. Participants gave feedback that the timing of challenges was extremely important to their participation: "whenever [challenges] were coming in at the right time, I always took the opportunity" (P49). Scheduling matches in the future would ensure convenient timing; participants were quick to recognize this, suggesting we allow them to "give a window of time for people to accept in (I want to play within the next 2 hours)" (P53), or "set a time when you are free [and] be able to reschedule" (P23). Therefore we added functionality that will allow participants to send challenges for specific times of the day or week. This feature could be critical to continued interest in the system in the future. One participant stated they would like to keep using the system, "if it was something where I had a little bit more control over the scheduling" (P54).

Along with the scheduling feature we found that expanding the type of challenges that participants can issue, could be a key improvement. We believe this feature would be received enthusiastically by participants, who expressed interest in issuing challenges for "other activities like soccer at the lab" (P27), or using it to "find someone to go to the gym with because I like the gym" (P21). Others thought it would be more interesting if "people could propose whatever challenge they wanted" (P25).

Opinions on the competitive aspect of the system were mixed. Some participants disliked having to compete, saying "when I realized I was out of the competition, it was disincentivizing" (P12), or even, "I hate competitive games" (P22). Others were highly motivated by competition. P63 said that, "one of the most fun things for me was competing with other people". He said, "I think the way you visualized a lot of the points earning was a good start", but wanted us to do even more to display his performance relative to everyone else in the competition. These differing opinions on competition highlight the importance of tailoring the intervention to individual preferences. In the future we could alter the system so that some participants can perform challenges without competing.

One of the major questions we have after analyzing the results of the study is whether it is appropriate to tackle both sedentary activity and social isolation at the same time. Our results appear to suggest that performing challenges does work to reduce sedentary activity, but many participants were deterred from participating because they were unwilling to meet strangers. Several participants wanted to "be able to challenge someone you know rather than a stranger" (P27), or even to do "individual activities" (P30). Although at least one participant said, "I like the random aspect of it" (P54), it is clear that many find it intimidating. To deal with this, P33 suggested allowing each user to create a group of friends within the system so that challenges from a stranger are directed to the group, and two people from the group go on the challenge together. She reasoned, "people are more comfortable meeting new people when they are with someone they know" (P33). Another suggestion was to "have the competition ask you to play with a specific person to get to the next round" (P13).

Despite there being room for improvement, participants were enthusiastic about the Challenge system. We received comments like, "Anything that forces people to connect in a busy place I think is valuable. I think it's really a nice idea" (P54), and that the challenges "went well" and were "good and cool" (P34). Several participants liked their tracker so much that they purchased their own. When participants were asked if they would like to keep using the system now that the competition is over, they responded, "I would love to and get more people using it" (P26), and "Yeah. I would like to do it again" (P46). These comments reveal that there is sufficient interest in the system to warrant further study. Considering our results show that performing challenges does have the intended effect of reducing sedentary activity, we believe further work on the system would be both beneficial to, and well received by, the Media Lab community. After making several of the improvements listed above — focusing on the timing and future scheduling of challenges, how to effectively contact participants, adding other types of activities, and making the aspect of meeting new people seem less formidable — we intend to take P53's advice to "extend the study longer" by performing a follow-up study.

4.3.5 Improvements Made Based on Initial Study

4.3.5.1 System Enhancements

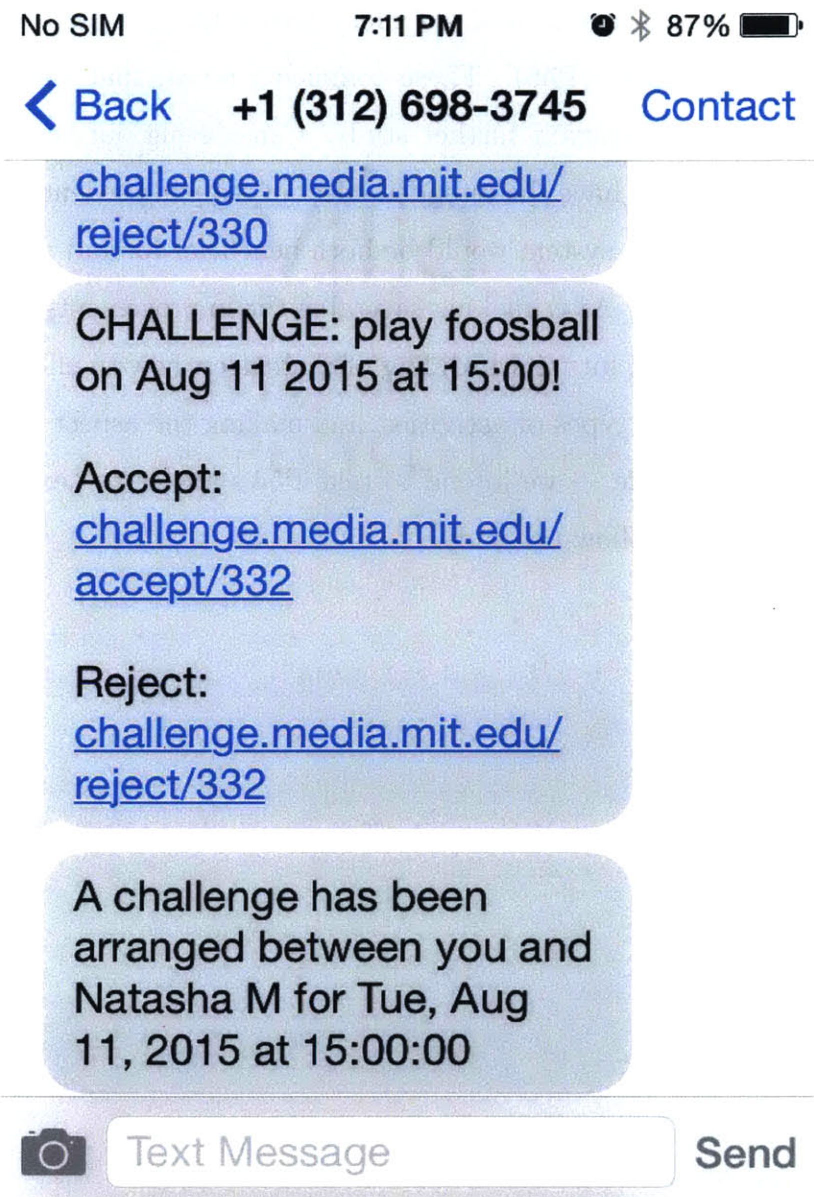


Figure 4-9: Users receive text messages on their phone when they are challenged to an activity, when the challenge is arranged, to remind them of a scheduled challenge, and if their partner cancels the challenge.

Given the feedback from participants, we added a number of new features to the Challenge system. Firstly, we added the ability to send challenges directly to participants' phones via SMS. Participants can choose to add their phone number to their account; if they do so, challenge request and arrangement emails will be sent via text message as well as by email. Figure 4-9 shows an example of a challenge text that was received on a user's phone as part of the study.

A second obvious improvement allowed participants to schedule challenges for times that are most convenient. Participants can select to book a challenge at the time and date of their choice, and the system will send out requests to other participants to see who is available during that time. As usual, once another participant accepts, the system will inform both users that the challenge has been arranged. Upcoming challenges can be viewed from a calendar on each user's home page (see Figure 4-10). Participants can use the calendar to click on scheduled challenges that they have requested and view information about them, even if no one else has accepted the challenge yet. They also have the option to cancel the challenge through this view. If a partner has accepted the challenge, clicking on the challenge will show the partner's name, photo, and contact information. In case users forget about an upcoming scheduled challenge, the system will send them a reminder about it a few minutes before it is scheduled to take place.

Not only do scheduled challenges make it easier for participants to make time to participate, it also takes the pressure off participants to respond to challenges quickly. When a scheduled challenge is issued, the system calculates the amount of time before the challenge will take place and divides by the number of users currently in that competition, in order to calculate the new expiry time of the challenge request. Therefore a user has much longer (sometimes 24 hours) to decide whether they will participate in a scheduled challenge before the request expires.

Allowing for scheduled challenges also enables the system to include more types of activities that need to be done outside of work, such as going for a run or playing tennis. After feedback from the initial study, we allowed participants in the second iteration of the study to vote on which activities they most wanted to be included, as part of the pre-study survey. Based on these results, we included six new activities in the second version of the system: running, swimming, going to the gym, tennis, rock climbing, and yoga. We also included the ability for administrators to add special

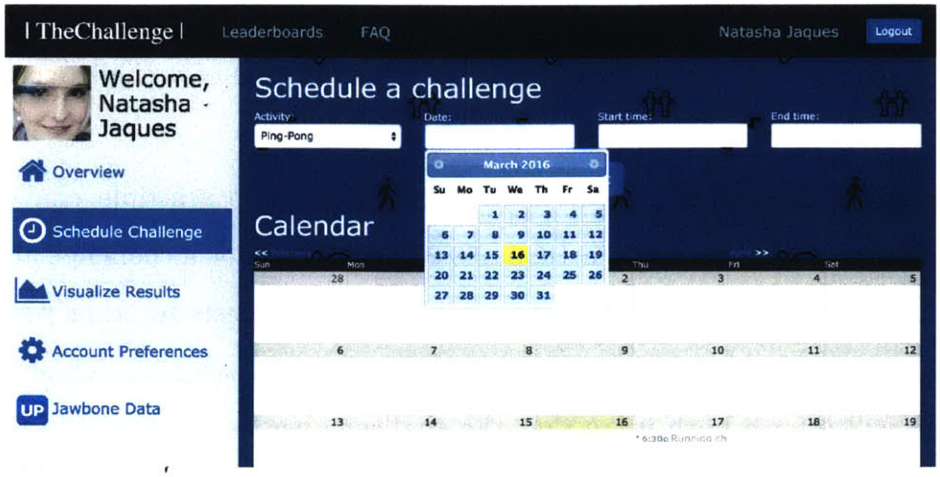


Figure 4-10: The re-designed app allows user to schedule upcoming challenges, view information about them in a calendar, and directly link their Jawbone Up Move device to their account.

weekly challenges: temporary challenges that would be worth extra points, but only last for a short while. These could consist of anything from 'take the stairs instead of the elevator', to 'build a snowman', or 'go sailing' - anything that seemed topical and engaging to participants. We decided to add this feature so that participants could propose any type of challenge they wanted, but before it was sent out system administrators to ensure it was appropriate could vet it.

The first study suffered from a significant amount of data shortage due to disengagement and attrition. As mentioned previously, 5 participants received a tracker but never used it or provided data. In order to be able to detect this type of behavior in the second round of the study, we integrated our application directly with the Jawbone API, allowing us to automatically pull participants' tracker data nightly. Previous to this, participants' data was obtained by asking them to download a .csv file from the Jawbone website, and was thus collected much less frequently. Obtaining the data automatically allowed us to add functionality to the administrator page of the site that would visually highlight any participants who were not wearing their tracker. We also used the Jawbone integration to automatically update participants' step and points counts on their user home page, in order to make the site more engaging.

Several other small features were added to improve the design. For example, the system itself automatically triggered approximately 3 challenges at random times during each work day, to ensure participants received sufficient triggers. The FAQ page was updated to more clearly explain Challenge mechanisms such as expiry time. We also added the ability for users to attach a personal message to their account, which was sent out along with their challenges. Although we wanted the challenges to remain anonymous so that users would be encouraged to play with new people and strengthen the community of the lab, several participants in the initial study mentioned that they found the idea of playing with strangers off-putting. Therefore we encouraged users to attach a personal message such as, "I'm a new student and would like to meet people in the lab". This message would be seen in any challenge originating from that user. We hoped this would make the challenges seem less intimidating.

Finally, based on expert advice, we also added a feature that sends the participants a weekly update email informing them of the current leaders in each competition and their points. The email then goes on to describe how many points the participant herself will need to reach third place in each competition. Because in some cases only a few challenges needed to be completed to reach this point, it was thought that these emails would encourage participation from users who had not previously engaged with the system. The competitive aspect of the system appeared to be highly motivating for some participants; for example, P63 stated, "one of the most fun things for me was competing with other people".

4.3.6 Experiment II

4.3.6.1 User Study

In addition to adding system features, we made several modifications to improve the experimental design of the study. The most obvious improvement was to use a randomized control design in the second iteration of the study. Although the within-subjects design of the present study allowed us to assess whether participants' behavior improved over their initial baseline, it did not allow us to determine whether external factors could be causing this behavior. For example, the effects of external influences like Members' week could not be fully evaluated. Therefore, we built a special version of the website for users belonging to the control group, so that they could still sign on and view their uploaded Jawbone data, but could not access any features of the site. The intent was to ensure any difference in outcome between the control group and the experimental group was due purely to participating in challenges, rather than simply viewing steps on the site.

Several participants in the initial study complained of not understanding the purpose of the study or how to use the system. In addition, many other participants never issued a challenge using the system. To address both these issues, in the second version of the study we held orientation sessions for participants in which we explained how to use the trackers and the site. Participants in the experimental group got additional information about challenges, what they mean, and how to issue and respond to them. We expected that explaining the features to the participants would ensure they felt comfortable issuing and accepting challenges.

In the initial study there was also no way to objectively determine whether participants' social isolation had increased or decreased. We had included a survey question that read, "How many different people from the lab do you spend time with on a weekly basis??? but we received feedback that this question was difficult for participants to answer, and the results were ambiguous. Therefore we added an additional questionnaire to the pre- and post- study surveys using the site itself. We asked participants to rate how well they knew each other user on the site, on a 5-point Likert scale that ranged from 'Stranger', to 'Best Friend'. We expect that this type of detailed data will allow us to better assess any change in the social community resulting from the study.

A final major change involved the monetary prizes awarded for gaining the most points in the competitions. Because a truly effective behavior change program would need to work and sustain interest in the absence of external monetary reward, we decided to remove all prizes. Participants could still earn points in the competitions and see the highest scorers on the leaderboards, but would not be rewarded by compensation based on those points.

MIT COUHES also approved the procedures for the second iteration of the study. Participants were recruited using posters hung up in the lab, which read,"Need a new gym buddy or running partner? Like trying out new devices? Want $70? Sign up for our new Challenge study!". In this iteration of the study, participants were compensated $2 for each day that they wore their tracker and contributed data. The study ran for five weeks, starting on August 17th and ending on September 21st.

The users were randomly partitioned into either the control group or the experimental group. However, the functionality of the Challenge system depends on having a large pool of users, which can respond to challenges. This effect is especially magnified in the new system given that there are now 9 competitions rather than 3, and users are likely to be more spread out among them. Therefore we strove to randomly partition twice as many participants into the experimental group versus the control group. Due to some participants dropping out before the study began, we were left with 17 participants in the control group and 37 in the experimental group, for a total of N=54 participants.

4.3.6.2 Results

Adherence to the study was improved in the second experiment, with participants wearing the tracker for an average of 27 days (SD = 10.87) out of a possible 35. However, there were still three participants who provided no tracker data whatsoever, and an additional three that provided less than 10 days of data. We did not include these participants in the analysis. One additional participant appeared to be an outlier; as Figure 4-11 shows, this participant had over 350,000 more total steps than the participant with the next highest step count. When we inquired into whether there was a malfunction with the participant's tracker, we discovered that he was a member of the track team practicing for an event. Due to the extreme nature of this data, we were forced to discard this participant from the analysis. In total, we were left with data from 47 participants.

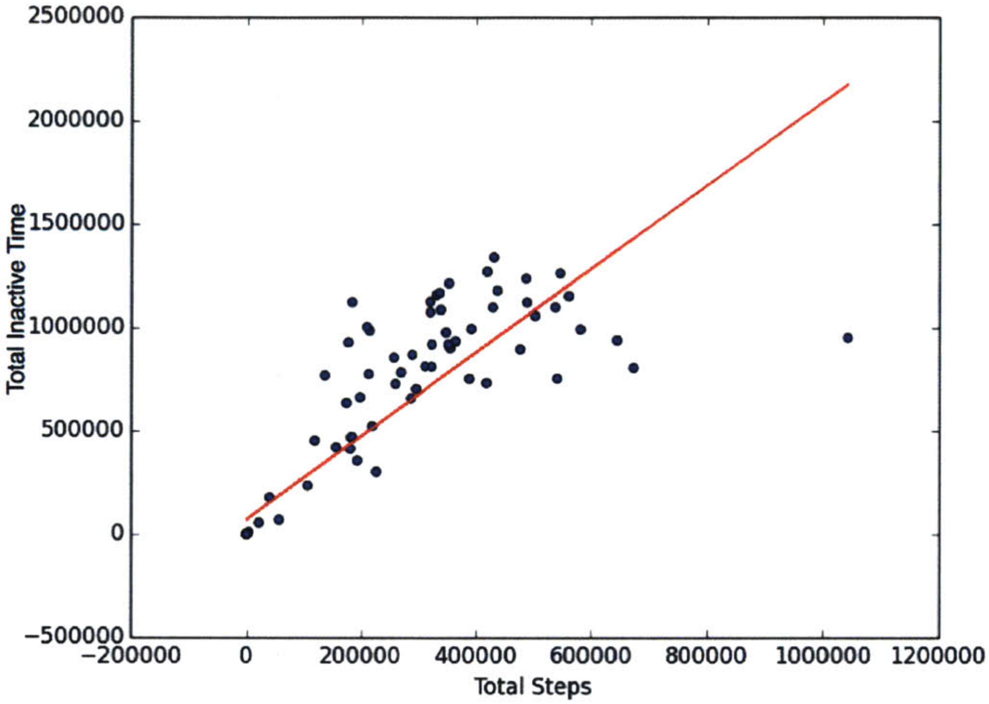


Figure 4-11: Each user's total steps and total inactive time. A clear outlier is present.

In addition to the users participating in the study officially, we also noticed an increased interest in the system from other members of the lab. We ensured it was possible for lab members that did not join the study to make accounts on the site, so that there would be a greater pool of challengers available to play. Figure 4-12 plots the total membership of the site in number of users from before the study started until well after it ended. As can be seen in the figure, membership continued to increase during this period as lab members presumably heard about the site and became interested. Note that users who joined the site but not the study were not given Jawbone Up trackers and thus could not gain points via their step count.

4.3.6.2.1 Activity results When all participants are considered, the control group and the experimental group did not differ significantly on measures of sedentary activity such as the longest time spent idle per day (see Figure 4-13a). Surprisingly however, the control group actually had significantly more steps per day than the experimental group, $t(34) = 5.21, p < .05$. Figure 4-13b shows the median number

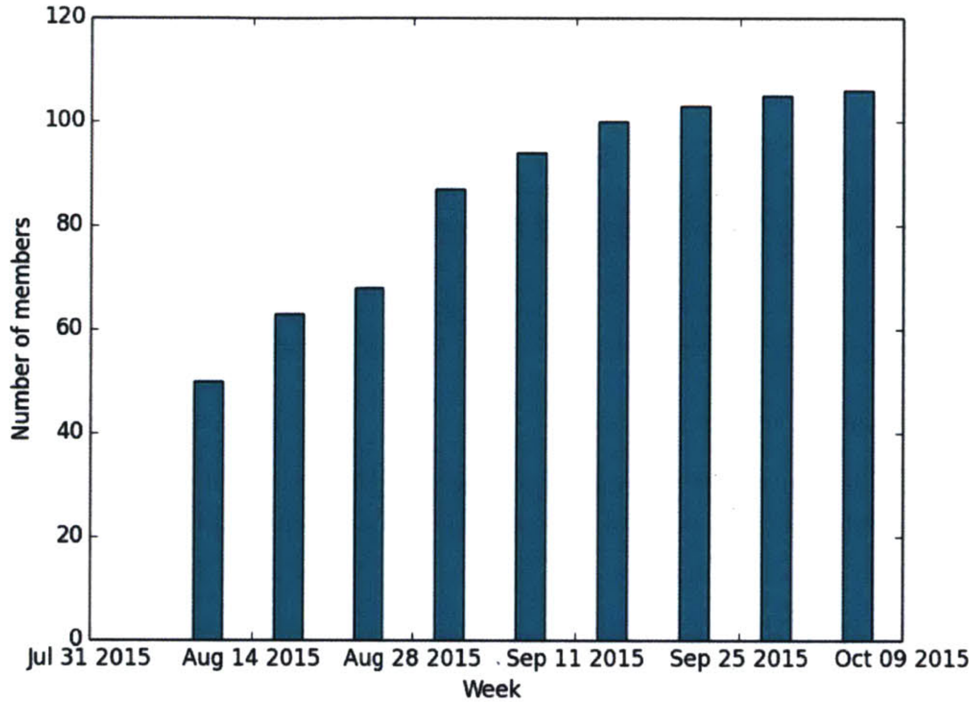
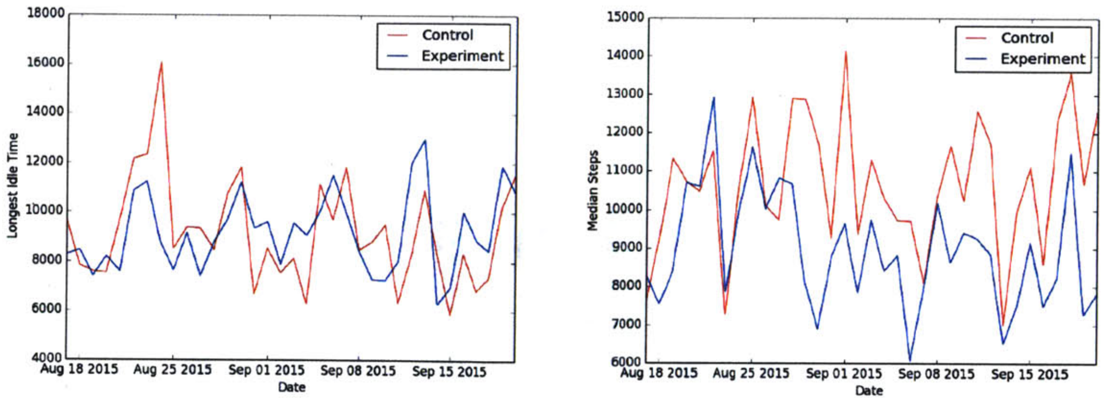


Figure 4-12: The total number of users on the site. The study began with 50 users on August 17th, and membership continued to increase well after the study ended on September 21st.

of steps completed for each day of the study by both the control and the experimental group. As is evident from the figure, the control group appears to consistently outperform the experimental group in terms of step count.

There are several possible explanations for this phenomenon. One is simply that the random assignment of participants into the control group actually produced a biased sample. In other words, perhaps the participants that were pre-disposed to have the highest step count happened to be assigned to the control group, and because the size of the control group was relatively small, this biased the results. We have some evidence to support this theory. We found that three participants with the highest daily step counts all belonged to the control group. Of these, we discovered that one was actually training for a marathon during the time of the study. Figure 4-14 shows the step performance of the control group versus both the experimental group from the second study, and the group of participants from the first iteration of the study. Note



(a) Longest idle time per day

(b) Steps per day

Figure 4-13: The control group and experimental group did not differ on measures of sedentary activity, but the control group actually had significantly higher daily steps.

that the group from the initial study is most similar to the experimental group, even before the intervention was applied. Since both the previous study and experimental group samples are larger, statistically we can expect them to better approximate the true population average. This theory is also supported by the fact that the control group had an average daily step count of 11,016 (SD = 5954) (compared to the experimental group's mean of 9402 ($SD = 4854$)). 11,000 steps per day is quite high, considering that studies of the typical step counts of American adults have shown the average to be closer to 6000 steps/day [87]

If the theory that the individuals assigned to the control group simply had a biased tendency towards higher step counts were true, we might still expect to see a higher relative change over time in the step count for the experimental group. In other words, we might expect that the intervention caused the experimental group to increase their step count more, above an initially lower baseline. Therefore we computed the average difference in step count between the first and last week for each participant, and compared the results of the control group to those of the experimental group. In this case, we found no significant difference between control and experimental groups, $t = 1.40, p = 0.17$, suggesting that the intervention did not cause the experimental group to become more active than the control group. Similarly, the two groups did not differ significantly in measures of change in sedentary activity between the first

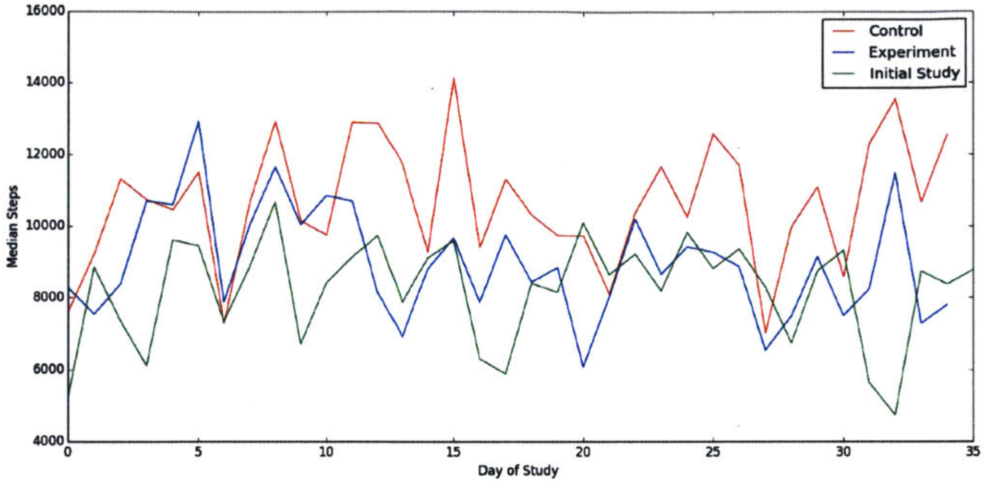
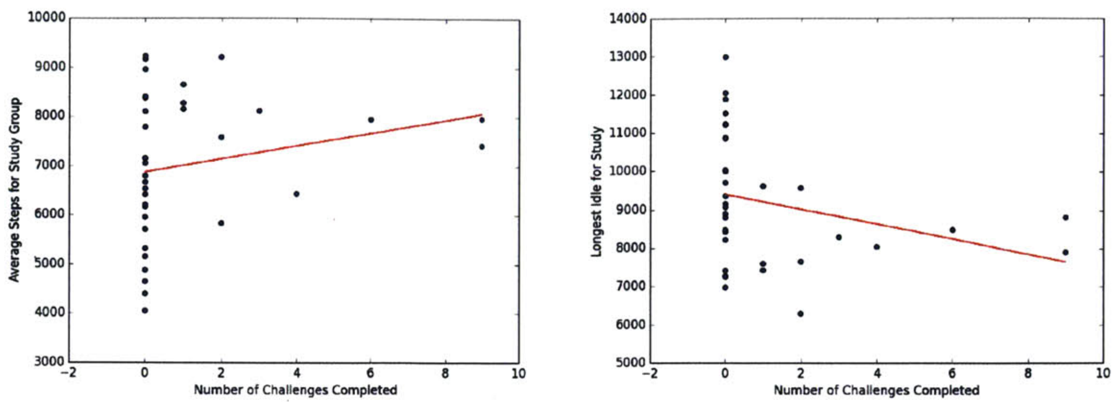


Figure 4-14: The step count of the previous study was more similar to that of the experimental group than the control group.

and last week of the study.

Due to these results, we may seek an alternative explanation for why the control group had significantly higher step count than the experimental group. Perhaps the Challenge intervention was actually de-motivating for participants, causing them to take fewer steps per day. Wearing a pedometer has been consistently shown to increase physical activity and step counts [78]. Perhaps the control group was strongly motivated by wearing the pedometer, but the experimental group was somehow "distracted" from this effect by completing challenges. However, we have reason to believe that completing challenges actually has a positive effect on both step count and sedentary activity. In the first study, completing more challenges was associated with higher levels of activity and steps, and lower levels of sedentary time. The same trends exist in this experiment. Figure 4-15 plots the number of challenges completed versus both step count, and the longest time spent idle in a day. Trend lines are fit to the data; we see that the trends are in the expected direction. However, the positive correlation between challenges completed and steps, $r(35) = .22, p = .21$, and the negative correlation between challenges completed and longest idle time, $r(35) = -.23, p = .09$, did not reach significance. The astute observer may immediately recognize why — there are a large number of points, which correspond to zero challenges completed.

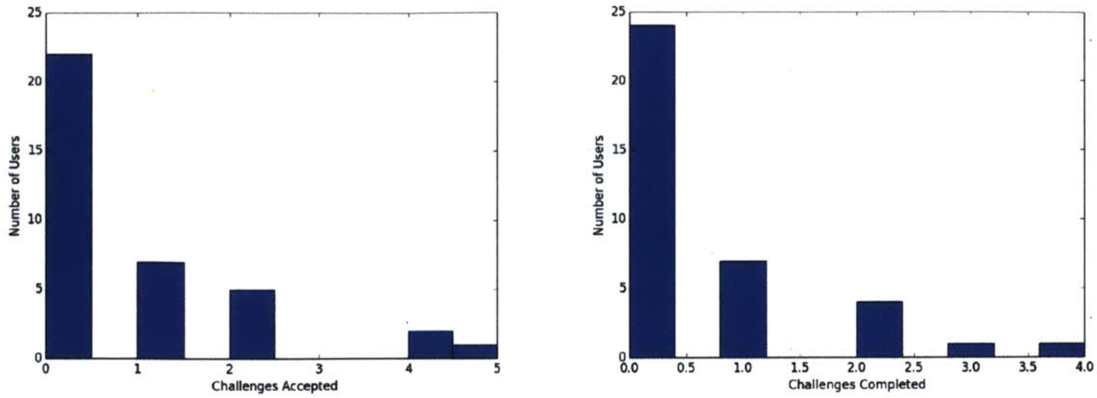


(a) Challenges completed v. step count (b) Challenges completed v. longest idle time

Figure 4-15: Trends exist in the data suggesting that completing more challenges is associated with higher step count and a lower longest idle time per day.

Upon further investigation, it becomes apparent that the problem is essentially one of participation. Figure 4-16 shows the number of challenges accepted and completed by each user. Over 60% of users did not complete even one challenge. Further, this iteration of the study does not contain the same type of super users as the previous one; the most challenges completed by a single participant in this study were four. Considering the study took place over five weeks, it is unlikely that such a small number of activities would have a significant effect on participants' activity outcomes. Discarding those participants who completed zero challenges and repeating the analysis verified this; however, the effects did not reach significance, possibly due to the small size of the remaining sample.

The question then becomes why participants did not complete a greater number of challenges. A first step is to assess whether the system worked as expected. Perhaps there was some malfunction in arranging challenges, or in the automatic challenge triggers. This does not appear to be the case. Figure 4-17 shows the number of challenges issued, accepted, and completed for each day. For each weekday, we see that at least three challenges were issued, demonstrating that the system was automatically sending out challenges as intended. Further, the number of challenges accepted very closely follows the number of challenges completed, suggesting that when users were willing to accept challenges, a challenge was arranged and they were able to complete



(a) Challenges accepted by user

(b) Challenges completed by user

Figure 4-16: The majority of participants did not accept or complete any challenges.

the challenge activity (this is also evident in Figure 4-16). The qualitative feedback we received, which will be discussed in Section 4.3.6.4, also assured us that the system functioned properly.

More closely examining Figure 4-17 gives us further insight into the problem. Note that on days when challenges were accepted and completed, there are high spikes in the number of challenges issued. This is because when a person has accepted a challenge and is waiting to play, the system will continue to issue challenges to other players in order to find a partner for that participant. The high number of challenges issued in order to find another player suggests that even when participants directly receive challenges, they are not willing to accept them. In fact, participants received an average of 15.56 challenges ($SD = 15.45$), but accepted an average of 0.88 challenges ($SD = 1.32$).

Why participants are unwilling to accept challenges will be examined further in Section 4.3.6.4, but for now it may be useful to note that there was a significant negative correlation between the number of challenges participants accepted and the longest time they spent continuously idle that day, $r(35) = -.34, p < .05$. When participants are spending less time continuously sitting idle, they also tend to accept more challenges. This correlation may be due to a third, external cause: participants' workload. Considering the population of graduate students participating in this study,

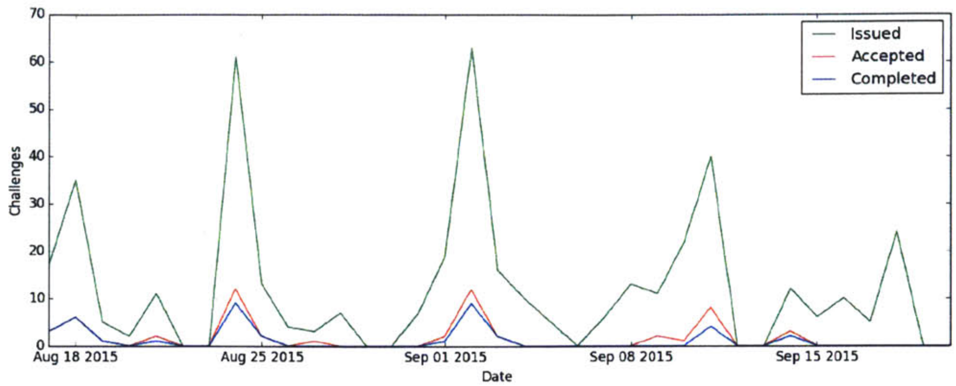


Figure 4-17: The step count of the previous study was more similar to that of the experimental group than the control group.

it seems likely that when they are under a heavy workload, they are both less willing to accept challenges, and more likely to spend long periods sitting continuously at their desk. Therefore, it would appear that any intervention designed to help participants break away from their desk briefly must be sufficiently enticing to overcome their desire to be productive; which in a competitive work environment, is no small feat. It seems that once the monetary prizes were removed in our study, participants no longer felt incentivized enough to participate in challenges.

4.3.6.3 Social network analysis

Although the lack of results for the second Challenge intervention is discouraging, there is still much that can be learned from the dataset that was collected. In addition to daily data about users' sedentary time and step counts, we also collected detailed social network graphs comprising all of the users in our study, both before the experiment took place, and after it ended. Figure 4-18 shows the post-study social network, computed from users' social ratings of each other. Recall that the ratings were collected on a 5-point Likert scale, with a rating of 1 corresponding to 'stranger', and a rating of 5 to 'best friend'. Figure 4-18 does not include ratings of 1, since these can be considered the equivalent of no social connection.

Although there were only 54 participants in the second study, Figure 4-18 has 125 nodes; this is because the graph includes those users of the site who did not actually

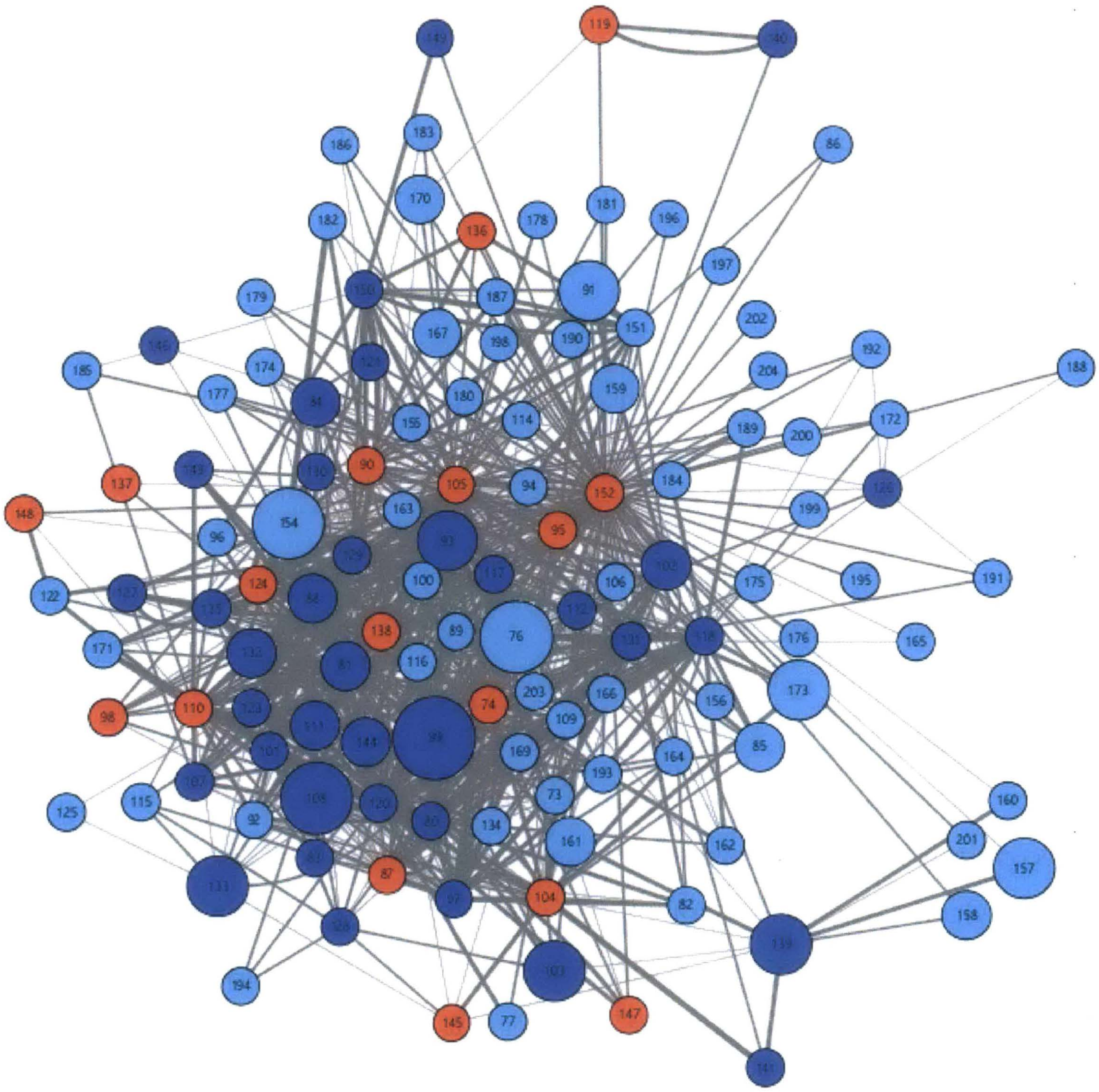


Figure 4-18: Post-study social network graph of all users of the Challenge system. Each node represents one user. Thicker edges represent stronger ratings; there is no edge if the person was rated as a 'stranger'. Dark blue users are in the experimental group, red users are in the control group, and the rest are users of the system who did not participate in the study. The size of the node is based on the number of challenges completed.

participate in the study and wear a tracker. It is still informative to include these users since, as we will see later, whether or not a participant's friends are also in the study has an effect on their participation. Further, there were a number of users who

	Rating				
	1: Stranger	2: We've met	3: Acquaintance	4: Friend	5: Best friend
Rating proportion	0.7162	0.0701	0.1066	0.0924	0.0146
Rating reciprocity	0.9025	0.5618	0.6116	0.7129	0.6848

Table 4.2: Proportions of ratings within the Challenge users (experimental and control) dataset, and the degree to which they are reciprocated.

did not participate in the study officially, but who were still active users of the site and completed several challenges. The size of the nodes in Figure 4-18 represents how many challenges were completed by each user. The overall density of the graph shown in Figure 4-18 is .097, the diameter is 2, and the average shorted path length is 5.03. Table 4.2 shows the proportion of various ratings within the dataset. Interestingly, a large proportion of ratings are 'Stranger', suggesting that many participants had never met each other. This could be due to the fact that in addition to a number of graduate students, the study contained a large cohort of visiting students and staff members who tended to only be acquainted with a small subset of the participants. The second row of Table 4.2 shows the number of ratings of each category that were reciprocated; for example, if 90% of 'Stranger' ratings were reciprocated, it means that out of all of the ratings in which user A rated user B a 'Stranger', in 90% of cases user B also rated user A a stranger. It is not surprising that ratings of 'Stranger' are strongly reciprocated, since it is fairly easy for people to tell when they have never met someone. What is more interesting is that ratings of 'Friend' are also highly reciprocated, suggesting there is a strong mutual understanding of when a relationship qualifies as 'Friend' status. Whether to assign someone a rating of 2 or 3 appears to be less clear.

The original intent of gathering the social rating data was to establish whether the Challenge intervention increased social connectivity significantly more for the experimental group than for the control group. However, as the previous section established, a large number of users did not actually complete challenges, preventing the intervention from having its intended effect. We therefore restrict our analysis to comparing the ratings of only those users that completed at least one challenge to the ratings of the control group[3]. Figure 4-19 plots the difference in the pre- and post-study ratings of these two groups. Since the dependent variable (rating) is ordinal and not continuous, it is inappropriate to use a statistical test like the Mixed Model GLM to analyze the data. Instead, we compute the difference between the pre-study and post-study rating for each pair of raters in the data. We then apply a Kruskal-Wallis test, to rank the differences and compare the rankings of the control group to the experimental group. The test revealed a significant and sizable difference in the rankings between the experimental group and the control group, $\chi^2 = 22.396, p < .001$; the experimental group's rankings ($M = 451.14$) were higher than those of the control ($M = 498.07$). This suggests that for those users that completed challenges, the intervention improved their social connectedness as intended.

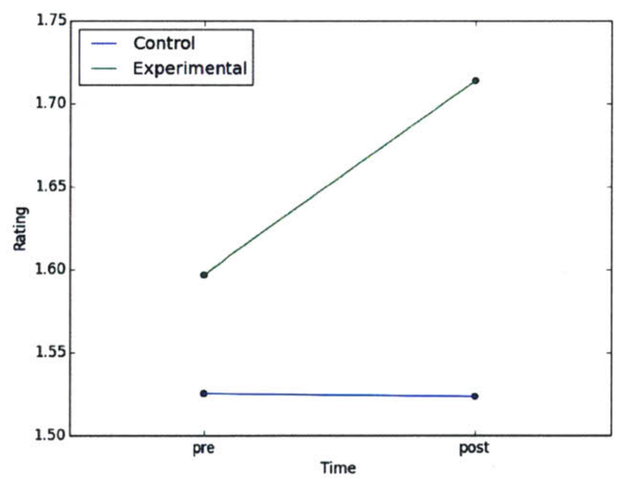
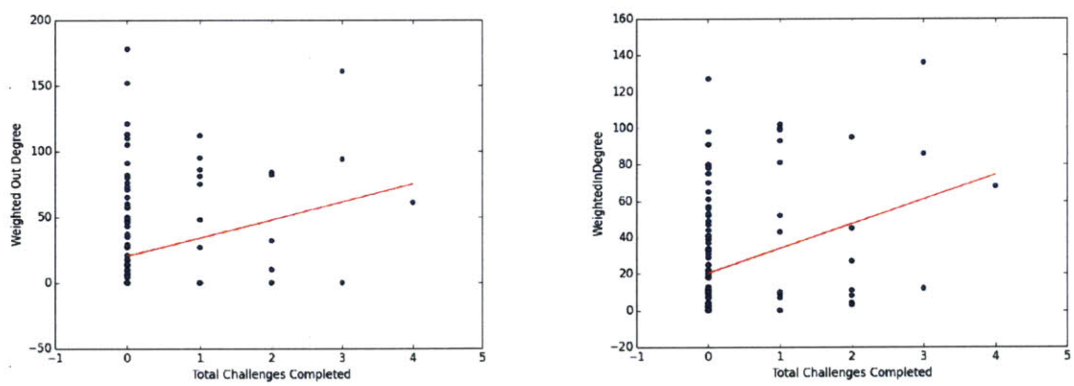


Figure 4-19: The change in average rating from the pre-study survey to the post-study survey for the control group, and the members of the experimental group which completed at least one challenge.

Interestingly, while completing challenges might improve users' social connectedness, their social connectedness might also be affecting their willingness to participate in challenges. Several correlations exist within the data between measures of social connectedness in the pre-study graph, and the number of challenges the

[3]When all users are included, the test is only marginally significant, $\chi^2 = 3.097, p = .078$.

users later completed. For example, the *weighted out-degree* and *weighted in-degree* of each node in the graph reflect the sum of the ratings (edges) outgoing from and incoming to that user (node), respectively. Both were significantly correlated to the number of challenges users completed; the effect size for the weighted in-degree, $r(125) = .304, p < .001$, was slightly larger than that of the weighted out-degree, $r(125) = .254, p < .01$. Figure 4-20 graphs these relationships. Users' pre-study *eigenvector centrality* — a measure which is higher not only if the node is more central in the network, but also if it is connected to other central nodes — was also related to the number of challenges completed, but this effect was only marginally significant, $r(125) = 0.175, p = .099$. Taken together, these results suggest that the social connectedness of the users at the beginning of the study may have an influence on whether they are willing to participate in challenges.



(a) Challenges completed by weighted out-degree

(b) Challenges completed by weighted in-degree

Figure 4-20: The degree of social connectedness of the users in the pre-study is related to the number of challenges they complete.

In fact, the users' place within the social network may have a strong effect on not only their willingness to complete challenges, but also their level of sedentary activity and their fitness behaviors. It is possible to partition a social network into *communities*, which are sets of highly-interconnected nodes. We partitioned the pre-study and post-study social network graphs using Blondel and colleagues' modularity based algorithm [88]. The modularity of a partition is a measure that compares the ratio of the density of links within each community to the density of links between the community and the rest of the graph. By applying Blondel's algorithm, we were able to discover four maximally modular communities within the pre-study network of users within the study (see Figure 4-21).

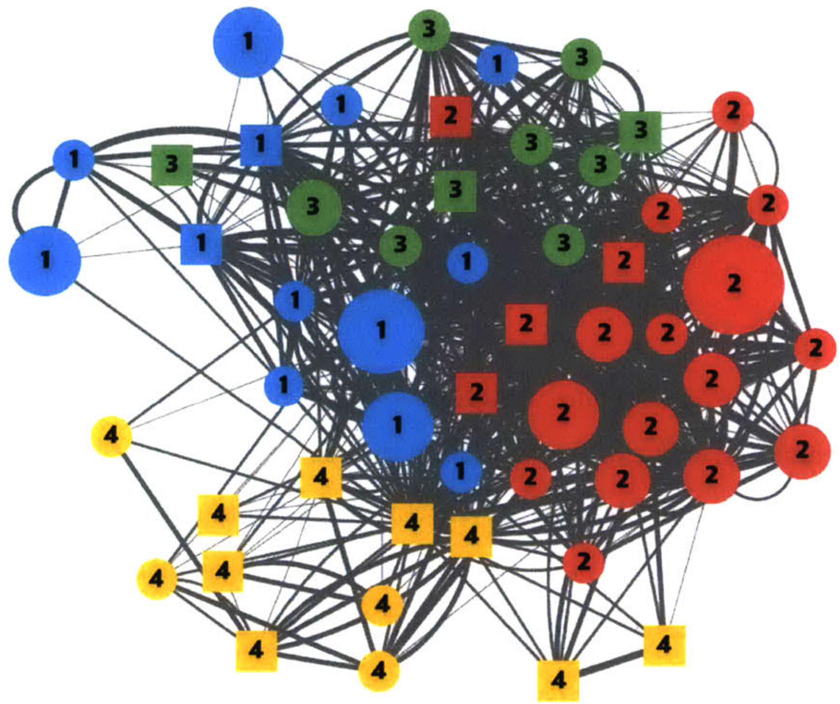


Figure 4-21: Communities detected in the post-study social network graph. Each community is labeled with its number, and assigned a different color. Shape indicates whether the participant is in the experimental group (circle) or the control group (square). A larger node size indicates that the user completed more challenges.

Considering the pre-study social network of only those users that participated in the study, we find some interesting trends. Table 4.3 gives several correlations that exist between a user's social community at the time of joining the study, and their later participation and physical activity[4]. The fact that the number of challenges accepted and completed are both significantly correlated with the user's social network community seems to suggest that users tend to participate in challenges as whole social communities; that is, if a user has friends who are participating in challenges, she may be more likely to participate as well. This effect is apparent in Figure 4-21, where the number of challenges completed by each user is visualized using the

[4]Note: making several significance tests at the .05 level may increase the risk of Type I error. We encourage the reader to examine the effect sizes to assess the robustness of these relationships.

	Correlation with community	
Measure	r	p
Total challenges accepted	**-0.333**	**.014**
Total challenges completed	**-0.334**	**.013**
Avg. daily longest idle	**0.276**	**.050**
Avg. daily longest active	0.245	.083
Longest idle difference	0.284	.093
Total inactive difference	**0.358**	**.032**

Table 4.3: The Pearson's r correlation between measures of participation and activity, and a user's pre-study social network community. Entries in bold are significant at the $\alpha = .05$ level.

node size. Community four has a high proportion of users in the control group, and it appears that none of the users within the study group in this community were interested in completing challenges. In contrast, community two had a few highly active users, which seem to have encouraged other users within that community to participate as well. Similar effects are visible in communities one and three.

As Table 4.3 demonstrates, communities of users are also similar in their sedentary activity and the amount of time they spend active. The *longest idle difference* and *total inactive difference* measures compute the change in the amount of time spent continuously idle and the total idle time between the first and last week of the study for each user. We see that these measures also appear to be related to social community membership. This suggests that a user's activity trajectory — whether they are improving or not — may also depend on their relationships.

These relationships exist not only with users officially participating in the study, but with other users of the system as well. The full social network including all users at the end of the study is worth investigating, because the users outside the study have an effect on the behavior of the actual participants. Figure 4-22 plots the five communities that were detected within the full post-study social network, which again relate not only to users' social connectedness, but to their participation and physical activity. Figure 4-23 analyzes these effects by illustrating how the communities differ on several measures of interest. Each measure is z-scored to enable comparison; each

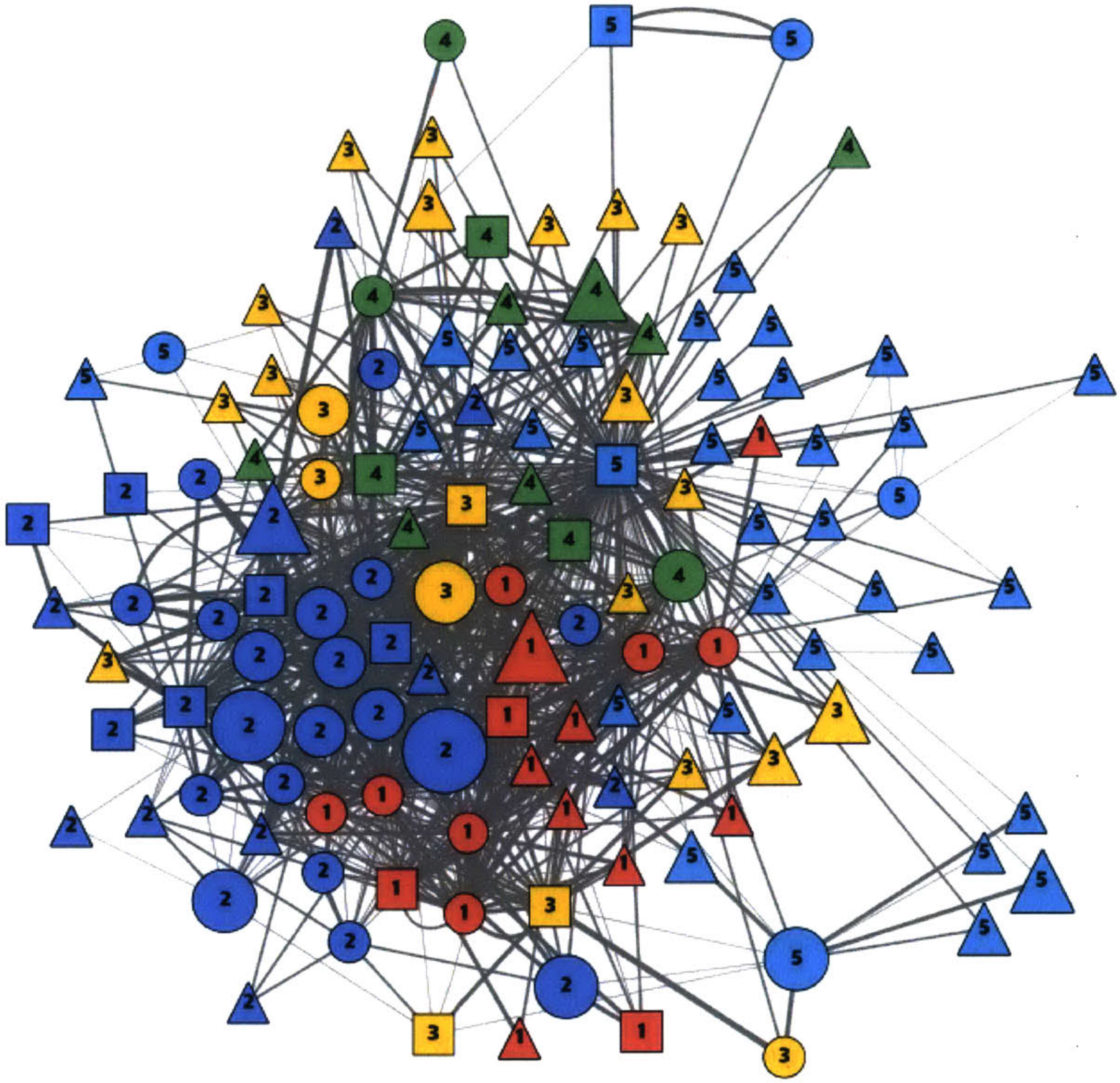


Figure 4-22: Communities detected in the post-study social network graph. As before, communities indicated with colour and numbered. Shape indicates whether the participant is in the experimental group (circle), control group (square), or a user that did not participate in the study (triangle). Larger nodes have completed more challenges.

has a mean of zero and a standard deviation of one (e.g. $z = \frac{x-\mu}{\sigma}$). Therefore if a bar in Figure 4-23 has a height above zero, then the community scored higher on that measure than the group average; if the height is below zero, the community scored lower than average. The error bars reflect the standard error ($\frac{\sigma}{\sqrt{n}}$).

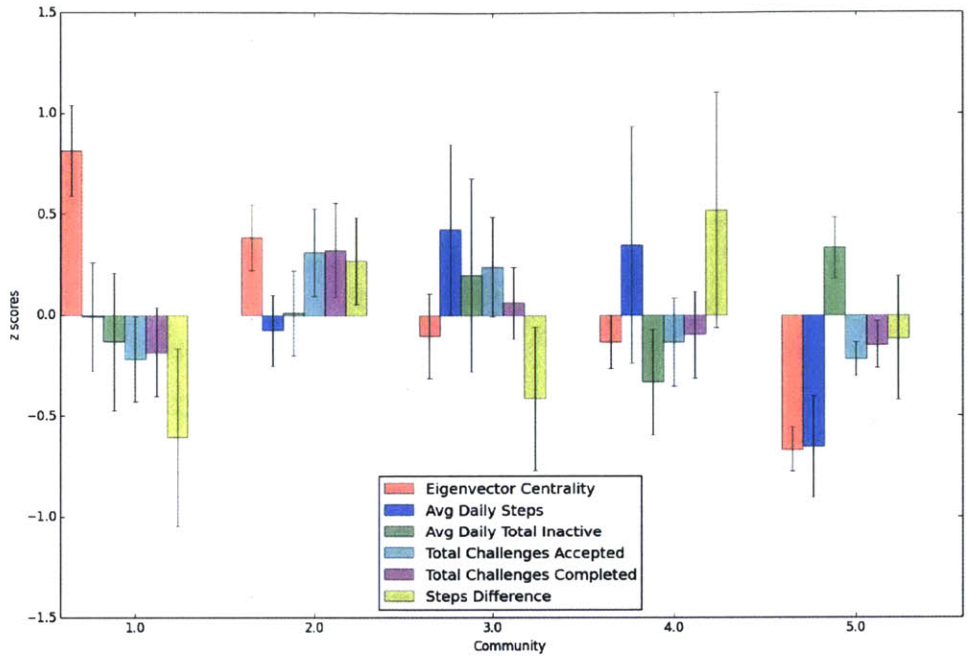


Figure 4-23: The social network communities differ on measures of participation and activity. The height of the bar indicates the z-scored value on a particular measure, and the error bars show the standard error.

Figure 4-23 reveals some interesting trends. First, we see that both community one and community two have much higher eigenvector centrality than the other communities, suggesting that these communities are more central in the network. This can easily be verified by examining Figure 4-22. However, while both community one and two are central, only community two appears to be actively participating in the challenge intervention; community two both accepted and completed a higher number of challenges. Once again we see that whether other members of a user's social community are completing challenges seems to strongly affect their willingness to participate. Figure 4-22 shows that community one had only one user, who was not in the study, completing challenges, but community two had a number of actively participating users.

In addition to completing more challenges, community two also had a higher step difference (see Figure 4-23). The step difference is the change in the amount of steps taken in the last week of the study from the amount of steps taken in the first week. If this measure is higher, it suggests that the participant's step count improved over the course of the study. These results may suggest that community two comprises ideal users; those that are actively participating, and also becoming more active.

Communities three and four appear to be relatively average; they contain a mix of users within the study and without, and score moderately on measures of participation and activity. An analysis of community five, however, reveals an intriguing hypothesis. Community five has one of the lowest proportions of users actually participating in the study (only 17%, whereas community one and two both have over 50% of users as participants in the study). In addition, community five has a much lower score on eigenvector centrality, meaning that its members tend to be at the outskirts of the social network. By examining Figure 4-23, we can see that community five also has a much lower average daily step count, a much higher average total inactive time, and a lower number of challenges accepted and completed. It may be that the users who are in the study but are a part of community five have the strongest social ties to people outside of the study, and therefore have the least incentive to participate in the Challenge system. They may also be disincentivized from even counting more steps with their tracker, since their friends are not wearing a tracker, and thus competing for step counts with their friends is not a salient motivation. Once again, it would appear that the users' social network has a strong influence on their participation and motivation.

One other insight worth noting is that the communities that are highly central in the network also tend to have a much higher proportion of participants that are actually in the study. Indeed, participants in the study have a significantly higher weighted in-degree than those who are not participating in the study, $t(53) = 5.078, p < .001$. This may seem like a trivial insight, since the participants within the study are the ones who are required to submit ratings, and therefore those outside of the study would not have rated each other. However, those within the study are required to rate all of the other users, so if the participants within the study were evenly sampled from all of the social groups within the Media Lab, then all users would have about the same amount of incoming ratings. Since this is not the case, it seems more likely that the study participants mostly come from a more tight-knit social community. No special efforts were made to recruit from within the same social network, but it seems likely that people would be more willing to participate if they knew several friends who were also participating. Since there is evidence within this study that a person's social network community affects their behavior and participation (as well as strong evidence in previous work, e.g. [89]), this may suggest that other studies which recruit participants from the same workplace (or university course) may suffer from biased results.

4.3.6.4 Qualitative analysis

The qualitative feedback provided by participants at the end of the study provides valuable insight into the functionality of the system, the lack of participation, and the effects of social pressure.

First, the feedback we received indicated that the system did work as expected. Participants confirmed that they were able to participate in challenges, and made comments like, "I would continue using the site. Great for finding ping pong buddies" (P132), and "I think it was well designed and fun" (P130). Nine users said that they met someone new as a result of using the system. P141 said that she, "became very good friends with a person", while P99 explained, "I met one new person, connected with friends, and also connected with one person I knew of but hadn't talked to before". Even those who didn't meet a new person did strengthen social connections; P129 stated, "[I] got to know people I did know better".

When asked if they would continue to use the system after the study ended, 44% said that they would, with one participant responding, "Yes, I love the idea" (P193). The fact that less than half of the participants are interested in using the app reflects our earlier findings; while some users appeared to benefit, a significant portion were not interested in participating.

The lack of interest may originate from two main concerns. The first is the desire to be productive. As outlined in Section 4.3.6.2.1, users appeared to be more willing to accept challenges when they have a lighter workload. The qualitative feedback added weight to this hypothesis, with users stating, "It was a super busy period that didn't allow me to take part in challenges" (p135), "If I had more time, I would be able to join challenges" (p80), and "Most of the time, I wasn't available to go do it" (p88). For some users, interest and motivation were not lacking, but work took precedence. When asked about continuing to use the system, P127 agreed, saying, "Yes, once my two hell-months are over, I think I'd like to pick this up for real". This feedback underlines the importance of making the challenges appear easy enough that users feel comfortable leaving work for a few minutes to participate. Finding times to challenge users when they are ready to leave their desk could also be beneficial. As P139 explains, "I participated in a few [challenges], mostly when the timing was good". As a future direction, we could investigate methods for detecting when users are most ready to participate, and sending challenges at that moment.

Perhaps the most controversial aspect of the entire study was the anonymous, random-assignment design of the challenges. Some users were enthusiastic about this aspect: "For me it's all about the social piece. I prefer interacting with new people than friends. I'd love it to remain anonymous" (P99). However, these users may have been in the minority. A considerable number of participants expressed their dissatisfaction with the uncertainty of not knowing who their challenge partner would be. "I don't like to meet new people for the time of a challenge and didn't have time. I would have preferred to challenge some people in particular instead of throwing a challenge and meet[ing] random people", explained P141. P118 was even more emphatic, stating that "the prospect of engaging in small talk with strangers 'just because' [was] very unappealing", and that "the energy required to interact with strangers for a short amount of time [...] outweighed the potential benefits". Several users explicitly stated that the thought of interacting with strangers prevented them from participating; P154 admitted, "I didn't prefer to initiate challenges for activities outside the Lab such as swimming, playing tennis, football etc., as I was hesitant to go out with someone new for the first time", while P140 said that he didn't participate in challenges "because I didn't want to do any of the activities with people I didn't know". Essentially, participating in the physical activities in the system while interacting with an unfamiliar person presents an unappealing prospect to many users.

In sum, while some users find the activities, competition, and socialization aspects of the system thrilling, the majority of users may find it intimidating. When users are willing to participate, they enjoy both physical and social benefits, but participation is low overall. A heavy workload and the anxiety associated with meeting strangers appear to be the two factors that most strongly discouraged participants. P132 may have summarized the situation perfectly when she wrote in her feedback, "I'm shy. Also, very busy :(".

4.3.7 Discussion

By eliminating the prize structure that was utilized in study one, participants no longer had an extrinsic force strengthening their motivation to practice healthy habits while at work. This elimination, however, shed light on other motivators.

It is clear that there are inherently at least three motivations in the average person; the motivation to be healthy, socially active and productive. However these motivators are not weighted equally. Since participants were advertised through flyers seeking individuals wanting to be more active, it can be assumed that the majority of the participants were motivated to be active in the work place. However, participation was significantly lower in the second round of the study. After post interviews, it was evident that productivity was a strong motivator to not participate in the system. This is also why there is currently an increase in sedentary time as people attempt to be as productive as possible and see breaks as a threat to their productivity. Our system attempted to have users spend time being active and take breaks from work and be socially active. Taking breaks, however, goes against the motivation to be productive. Although users typically want to be socially active, they want to be socially active in their social networks. Our system did not guarantee that they would be active with someone they knew which caused anxiety about having to meet someone unfamiliar. The initial system design was able to combat this by increasing the motivation to participate in the system by offering the possibility to win prizes. Without this prize system, the only reason to utilize the system would be a strong desire to be active that outweighed the desire to be productive.

When social networks are evaluated, it is clear that the motivation to be social with peers can outweigh the motivation to be productive. Those that were in social clusters that were more active were themselves more active and utilized the system more than those that were in social clusters that weren't as active. This leads to the conclusion that if users are placed in groups where their peers are high use users, they will also use the system heavily. A future iteration of the system could remove the anonymity and allow users to see whom they are challenging and who has challenged them. This, unfortunately, does not solve the productivity issue for users in social circles with low use users. Further research can be done to determine how to form mixed social groups so that low use users are mixed with high use users in the social cluster.

4.3.8 Study Conclusion

While the results of the first study were positive, the design of the system led to only short-term effects. In the second iteration of the system, the prize structure was removed to try to produce a system that could be used for long-term effects, since there was no conclusion to the competition. A control group was also added to ensure that the effects seen were due to the system and not extrinsic factors. However, these changes led to a striking decrease in the use of the system. After looking more closely at the social networks that were present, it was clear that there were social clusters that were more engaged in the system. Within these clusters there were similar effects to those seen in study 1. This reaffirms that this system, when used as intended, is effective at increasing active time, decreasing sedentary time and improving social cohesiveness in a closed community. The difference in structural components of the triggers used it both iterations of the The Challenge system are summarized in Table 4.4. Future work can be done to determine the best ways to encourage the use of the system by taking advantage of other inherent motivations.

	Motivation	Attention	Timing	Message Content
Study 1	√	√	√	√
Study 2	X	√	√	√

Table 4.4: Summary of Trigger Usage for Challenge System

4.4 Conclusion

When adapting existing technology in the persuasive system design process, these devices can be associated with a desired behavior. This allows a single device to take on multiple uses and can help motivate a user by making a desired behavior easier to do and track. However, there exist several motivations that are weighted differently for each user. Productivity is commonly one of the strongest motivators. Behaviors that threaten productivity need to be reinforced heavily to overcome this. Enforcement varies by user, but social influence and gamification appear to be effective means to reinforce users' desire to improve behaviors related to well-being. By using prizes and social influence we were able to reinforce participants motivation to take breaks from work in The Challenges system. However, without prizes social circles need to be thoughtfully planned out. However when productivity is not a factor as in the case with Smile Catcher, participants are sufficiently motivated with a simple game structure in which they are rewarded in points. When designing systems it seems to be important to consider motivators are being threatened by the new behavior and what can be done to either avoid competing with this motivation or to strengthen other motivators.

Chapter 5

Attracting the User's Attention

5.1 Introduction

When designing persuasive systems for well-being it is imperative that nudges are seen and processed by the user at the moment the desired behavior is intended to occur. Head-worn devices, such as Google Glass, have the advantage of containing a screen that is easily seen by the wearer at all times, in contrast with other device screens, which can be hidden in pockets or simply easily ignored. Because of this, these devices provide a great platform to develop and test applications that require just-in-time information and interactions based on the user's current context [90]. However, it cannot be assumed that all messages presented to a user are seen simply because the display is at eye level. Therefore, it is important to experiment with situations where a user is distracted and test how the system can attract the users's attention in these situations.

Mobile devices have become increasingly efficient at diverting user attention and interrupting daily activities, so much so that many attempts have been made to make these devices more contextually aware [91][92]. In related work researchers have shown what methods are effective in grabbing a user's attention on a mobile device [93]. However, not much is known about which methods work best to attract the users' attention with head-worn displays. Ideally the messages will be minimally disruptive, unlike notifications given through mobile devices. It is tempting to apply the same attention grabbing principles to head-worn displays, but with simple user testing effective techniques can be determined for a particular device. This chapter describes two user studies that were conducted to test if a user engaged in certain tasks may miss messages that are presented on a head worn display and which method is most effective in attracting the user's attention. The findings may be relevant to new form factors for head-worn transparent displays that are emerging, such as Microsoft Hololens, Sony smart eyes glasses, etc.

5.2 Pilot Study 1

Ten users were asked to perform three activities. Each activity was meant to simulate a common task, which a user may be engaged in (computer interaction, conversation and physical activity such as walking). Each task required visual, auditory, or spatial focus to complete the task. During each of these activities a word was displayed on the Google Glass screen accompanied by one of three interventions: (1) auditory cue (a beep can be heard prior to displaying the message), (2) visual cue (the screen flashes prior to displaying the message), or (3) no cue. The Google Glass uses a high-resolution display equivalent to a 25-inch high definition screen seen from eight feet away [94]. However, the display is not directly in line of sight, but slightly above the right eye. Each of these cue conditions was tested during each of the three activities for each of the ten test subjects. This pilot study evaluated which activities are the most distracting and which methods of delivering the just-in-time notifications were most effective and preferred by users.

5.2.1 Procedure

Each participant was asked to perform three activities. During each of the activities, three different words were displayed with a 15 second delay in between words. The participants were not told the words ahead of time and nothing besides the words was presented on the display. Each word was displayed with or without a cue as described in the introduction.

For the first activity, each participant was introduced to a computer game called "FallDown". The game, which was played on a laptop, consists of using the arrow keys to guide a ball down the screen without being trapped as obstacles advance in the opposite direction. The nature of the game required motor control, prospective memory and a high level of visual attention. The game had no music or sound effects that could distract from the auditory cues used during the experiment. The participant was asked to continue playing until notified to stop by the prompter, which required participants to start a new game when they "died".

During the second activity, participants listened to a series of 8 numbers (single or 2-digit) read aloud and were asked to repeat the series backwards. This continued until all three words were presented. This tasked required a high level of auditory attention and semantic memory, but a low level of visual attention.

The final task involved each participant walking back and forth while touching their nose with the hand opposite of the leg that is currently off the ground. The participant was asked to continue walking until enough time had passed for all three words to be presented. This task required motor control, moderate motor attention, perception (to walk in a straight line to avoid objects), and moderate visual attention.

Each task lasted about 1 minute. After each activity, the participant was asked how many of the words they observed by correctly reciting each word. They were also asked which of the cues they found most effective and how they would improve their cue of choice. Additionally, the participants were asked for any additional feedback they could give about the cues. Each participant was asked which method they preferred and if they would improve or change this method in any way.

5.2.2 Results

Figure 5-1 shows the average number of participants that were successful with word recall regardless of the cue type. We can see that participants had a significant lower recall rate when performing task 2, repeating number sequences. This led us to use task 2 as the distraction task in the second pilot study.

Figure 5-2 shows the effect of cue type regardless of the task. While not statis-

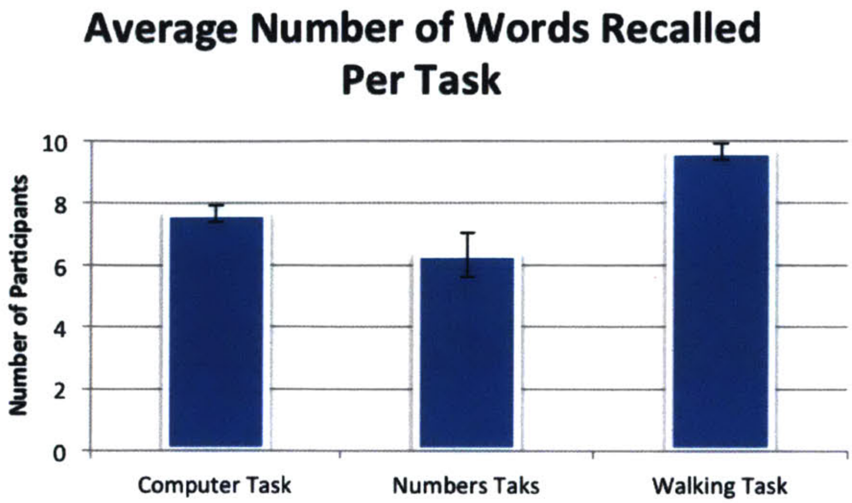


Figure 5-1: Number of participants able to recall the given words regardless of cue type (n=10)

tically significant, the data suggests that audio cues are more effective at grabbing attention.

Table 5.1 summarizes the results from pilot study 1. The number recitation task was the most distracting situation for our participants. During this task, the auditory cue was most effective at regaining the participant's attention. It is interesting to note that this was the only activity that did not require visual attention; yet visual cues did not outperform other cue types. The walking task was the least distracting and had equivalent performance across all cue types. Although the tasks were presented in the same order for each participant, we do not feel the high performance during the walking task could be attributed to learned effect or there would be upward trend across the three tasks.

Generally, the auditory cue outperformed not using a cue by 73% and the visual cue by 3%. However, the performance performance between visual and auditory cues was starker when user preference was taken into account.

Figure 5-3 illustrates the cue method preferred. Participants, at a drastically higher rate than any other cue or no-cue method, preferred auditory cues for all activities.

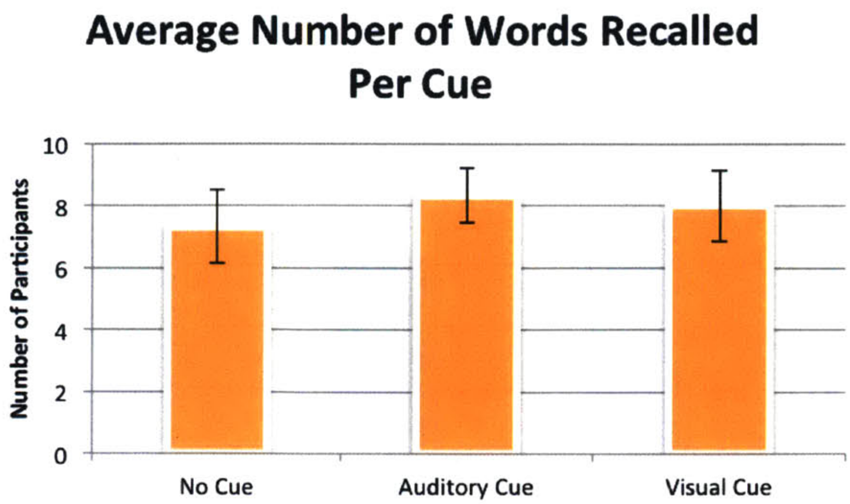


Figure 5-2: Number of participants able to recall the given words by cue type regardless of activity type (n=10)

	None	Audio	Visual
Task 1 - Game play	8	7	8
Task 2 - Number sequences	5	8	6
Task 3 - Walking	9	10	10

Table 5.1: Total number of participants able to recall the word given for each test scenario

5.2.3 Discussion

The information in Figure 5-1 and Figure 5-2 indicates that auditory cues are preferable for displaying just-in-time messages on head-worn displays. Intuitively, visual cues would be assumed to be more effective at grabbing the attention of wearers of head-worn displays intended for everyday use. Visual cueing creates a spotlight, which should guide attention faster to the exact location of where information will be presented. However, during particularly distracting tasks, wearers tend to look in off-screen directions. This could help explain why auditory cues work better in these situations.

Participants also found the visual cues to be "jarring", producing a jerking pull to attention. This goes against the common perception of cues used for mobile devices where "auditory cues are more public and intrusive in nature while tactile and visual

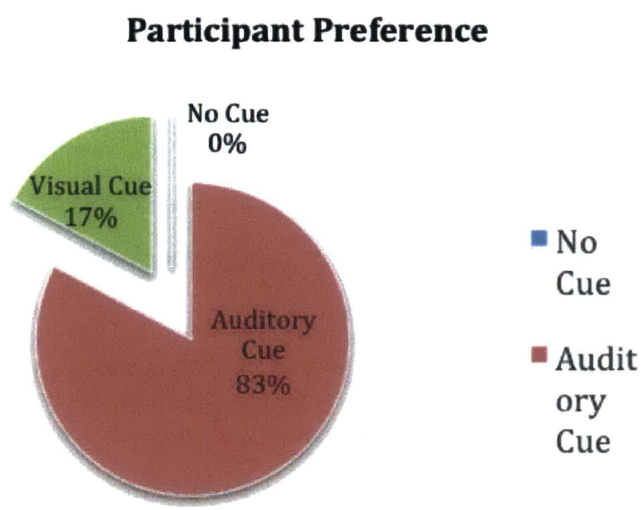


Figure 5-3: Average number of participants able to recall the given words by cue type regardless of activity type

cues may be more subtle and private" [93]. The most common complaint of auditory cues was the inability to be heard in noisy surroundings. In these situations it may be better to use a visual cue, however this pilot study focused on normal activities and situations.

Each of the tasks given to the participants used limited stimuli. All participants were familiar with the environment in which the pilot study took place and there were no noticeable visual or auditory cues besides the ones given. We would assume, due to the limited stimuli, that all of the intentional cues given during the pilot study would be perceived as relevant and be processed [8]. The decreased performance during the second task suggests that effective cueing is necessary for some tasks to ensure just-in-time messages are noticed and processed.

5.3 Pilot Study 2

Analysis of the results from pilot study 1 showed that the number series reversal task was the most distracting, and suggested that the auditory cue is better at getting the user's attention than the other two modalities. To test this further we ran a second pilot study to assess if auditory cues are better than visual cues at grabbing the user's attention.

5.3.1 Procedure

In the second pilot study, 30 subjects were asked to perform the number series reversal task while wearing Glass. During the activity five words were presented, each preceded by either a visual or auditory cue (identical to the cues from the first pilot study), where each subject was only exposed to one type of cue — assigned randomly. After all five words were displayed the user was asked how many words they thought were displayed in total and which words they could recall.

5.3.2 Results

Based on the results from pilot study 1, we wanted to verify that auditory cues would be more effective in situations that required focused attention, since these tasks have a higher chance of messages being displayed and not processed. From Figure 5-4 we can see that when using an auditory cue, significantly more words are remembered than when visual cues are used $p = .043$. Subjects belonging to the auditory cue group notice nearly twice as many words on average.

5.3.3 Discussion

From Figure 5-4 it is clear that participants who received the audio cues outperformed or did as well as participants given the visual cue. Figure 5-5 highlights that auditory cues also outperformed visual cues when participants were asked how many words they thought they saw. The average number of words that students thought were presented, regardless of whether they could remember each word was 3.23 for participants given audio cues versus 2.4 for participants given visual cues. While this result was not statistically significant it shows that participants saw more words than they could remember. This suggests that audio cues outperform visual cues when attempting to attract attention to a message on a heads-up display.

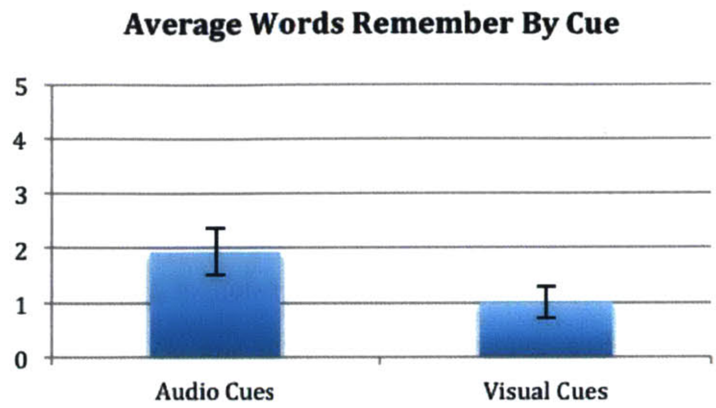


Figure 5-4: Words remembered by cue

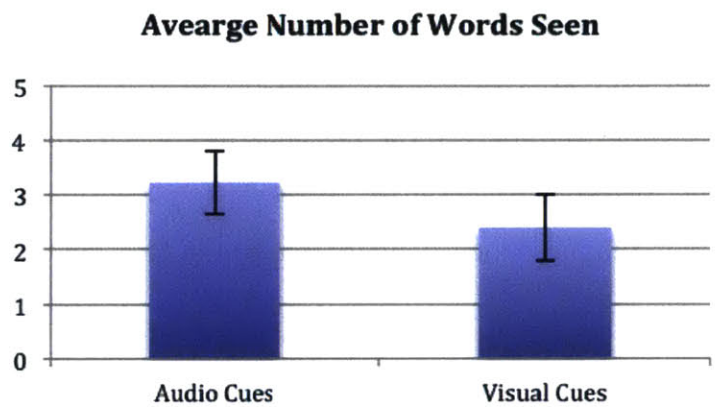


Figure 5-5: Average number of words participants thought they saw by cue type

5.4 General Discussion

Development for this pilot study was done using Wearscript [95]. The software causes a slight flash of the screen from black to white when cards are added. This flash is significantly shorter than the intentional flash and only occurs once. If this effect aided in attention grabbing we would not expect a significant increase in performance when using the auditory cue that we observed.

Future work could investigate the use of multi-modal cues. As [91] notes, for mobile cues it is advantageous "to avoid social misinterpretations as well as the problem of attention overload, it is desirable to design notification cues which combine the qualities of being subtle and public" [91]. Other future work could dynamically vary the cue used based on the task of the user as well as the context. For example, in a noisy context, a visual cue may be chosen.

5.5 Conclusion

In this chapter different ways to attract the attention of a user to a head-worn display were investigated. Previous work has shown that typical mobile technologies "cause resource depletion, which in turn restricts interaction with mobile devices" [96], whereas head-worn displays have the ability to decrease cognitive load by reducing attention-switches. However, as these displays become more common for everyday activities, it becomes necessary to find the most effective ways to display just-in-time messages to users for context-triggered attention-switches. For wrist or waist worn wearables, it makes sense to duplicate the methods used for mobile devices, as they have almost become wearables themselves. However for head-worn displays, visual cues may be viewed as the more logical choice due to the location of the screen. This pilot study has shown that auditory cues may be a more effective design choice in common distracting situations. Further work can be done to determine the best way to use both cue methods and when to switch between visual and auditory cues to ensure fluid interactions for the user without producing harsh alerts. When choosing technology for behavior change, it is important to ensure that nudges can be seen and processed by a user at the right moment. If a user does not notice the persuasive content, the probability of behavior change is low. Simple testing can be done to understand when a user is highly distracted. The designer can then choose to avoid presenting messages in these situations or determine methods that are effective at regaining the user's attention.

Chapter 6

Experiments in Varying Timing of Triggers

6.1 Introduction

Appropriate timing is important factor when designing persuasive systems. This chapter describes a system to persuade users to take breaks and be active by presenting messages on Google Glass. The timing of the messages is tested in order to determine how it can best be optimized.

6.2 Move Your Glass

Move Your Glass is a system that attempts to reduce sedentary time with notifications displayed on Google Glass that encourage the user to get up and move. Glass provides a tri-axial accelerometer similar to that used in current mobile phones. With Glass, there is the added benefit of consistent positioning on the user, whereas a phone may shift between different on-body locations depending on clothing, allowing Glass to collect more dependable data. The eye-level display on Glass also provides increased accessibility to transmit activity data back to users compared to current devices. This immediate accessibility can be used to convey feedback to promote healthier habits. Given these features, Google Glass represents an ideal platform for an activity monitoring and intervention application. This system was used as a platform to test if the timing of the message, with respect to the user's current context, makes a difference in producing a change in behavior. Specifically, we test whether reminding users at periodic intervals to take breaks from sitting or only displaying the message when the user has prolonged sedentary time is more effective. Table 6.1 illustrates what aspects of trigger structure are present in the Move Your Glass study.

	Motivation	Attention	Timing	Message Content
control	√	√	X	√
non-control	√	√	√	√

Table 6.1: Summary of trigger usage for Move Your Glass system

6.2.1 Related Work

Activity monitoring has previously been performed using a variety of wearable sensors and mobile devices. For each study, the optimal descriptors, elements that can be used to identify the activity state, derived from sensor data depended on the configuration of the sensors and the learning library used for activity identification. These works provided a starting point for determining the descriptors considered in this work.

Mobile phones have been widely used for activity analysis either as standalone devices or coupled with other sensors. For instance, Győrbnó et al. [97] used a smartphone as a central hub for logging data from wrist-mounted wearable sensors containing an accelerometer, a magnetometer, and a gyroscope. The intensity of the acceleration at each location was determined by loading all sensor data into a designated computer. Measured intensities were then compared to training data for resting, typing, gesticulating, walking, running, and biking using neural networks. Similarly, Kwapisz et al. [98] used data from tri-axial accelerometers on Android-based cell phones, sampled over 10-second intervals, to characterize six activities: walking, running, sitting, standing, and climbing up and down stairs. For each channel of the accelerometer signal, the average, standard deviation, average absolute difference, time between peaks, and binned distribution were determined. The overall acceleration magnitude was also used as a descriptor for each motion. These classifiers were used as training data for learning routines based on decision trees, logistic regression, and multilayer neural networks.

There have been several strategies developed to reduce sedentary time and/or increase physical activity. A review of such strategies found that the most promising focused on reducing sedentary time, treating it as independent of physical activity. It was also noted that the most effective techniques were self-regulatory, such as self-monitoring, problem solving and information on health consequences [99]. Therefore, it is no surprise that most devices developed to help reduce sedentary behavior focus on monitoring the user's current activity. However, a summary of activity is typically not enough to bring about behavior change. This chapter tests whether triggers given at a specific interval of time versus triggers given based on current sedentary behavior are more effective to invoke a change.

6.2.2 System Design

The Move Your Glass system is designed to run on a standard Google Glass device. Data from the tri-axial accelerometer on the Glass was sampled at 10 Hz (period of 0.1 s). The axes corresponding to the directions of Glass motion are shown in Figure 6-1. Sample raw data from the accelerometer can be seen in Figure 6-2.

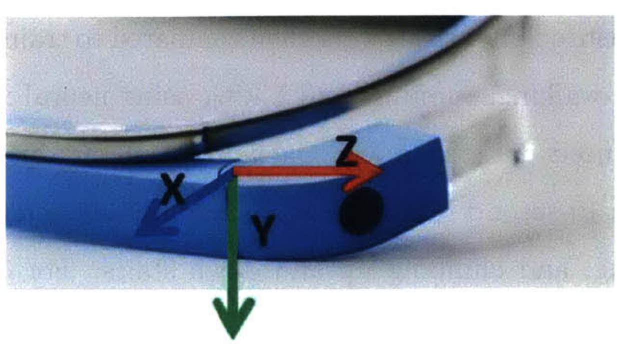


Figure 6-1: Coordinate system for the accelerometer on Google Glass.

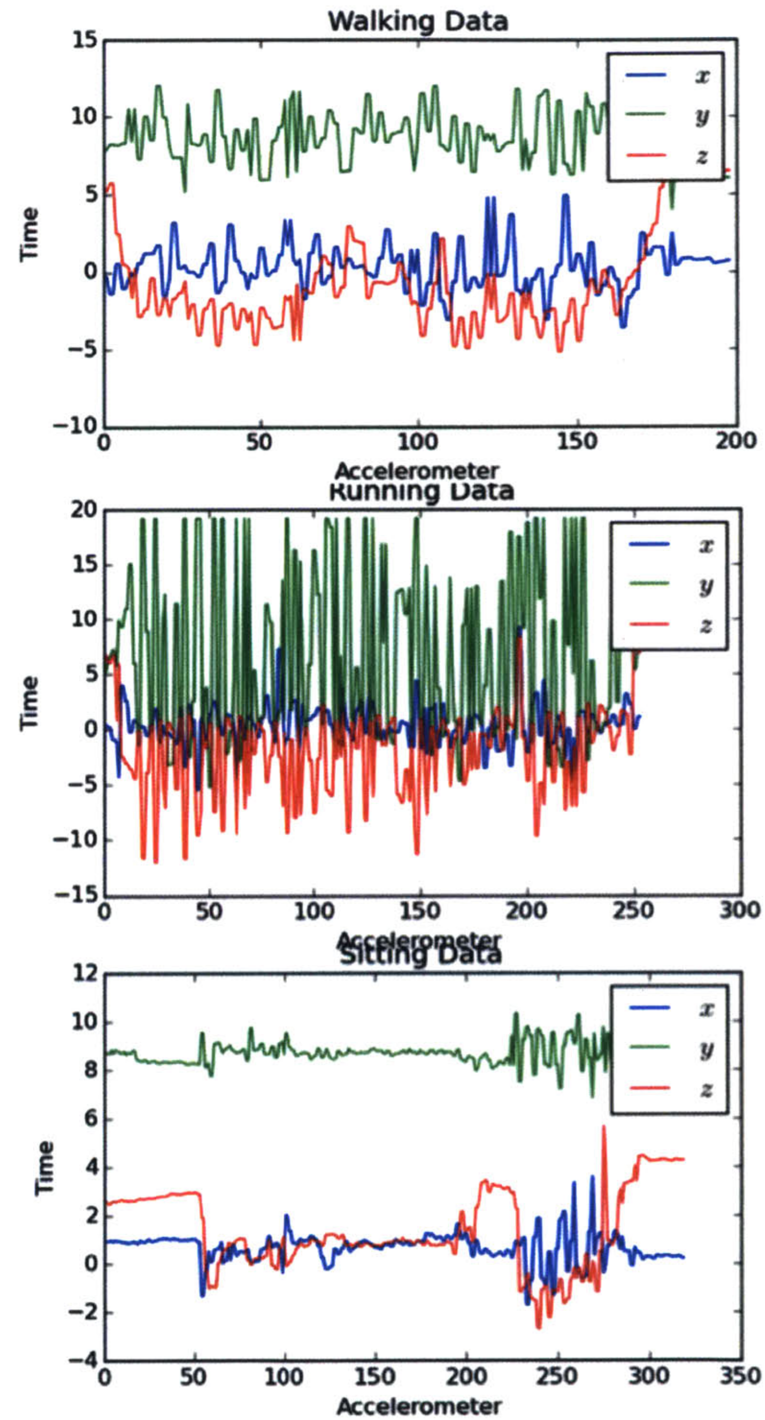


Figure 6-2: Raw accelerometer data during walking, running, and sitting. Gravity is included in the y-component of the accelerations measured.

6.2.2.1 Analysis

Based on the results of previous accelerometer studies, the following descriptive parameters are calculated from the accelerometer data:

- Mean acceleration in each direction
- Standard deviation in each direction
- Mean resultant acceleration magnitude (average of $\sqrt{x^2 + y^2 + z^2}$)
- Mean absolute difference in each direction (e.g. mean of $|x - \bar{x}|$)

These quantities are then used as the descriptors for the machine-learning algorithm. For accelerometer analysis, one user supplied preliminary training data for five instances each of walking and sitting performed. A queue of accelerometer points of length 25 (~25 seconds) is kept to predict the current activity based on training data. Activity recognition is performed using the k-nearest neighbors learning algorithm implemented in OpenCV [5]. K-nearest neighbors compares a given data input to its k closest matches in the training set and assigns a class based on the type of these matches. For each call of the activity recognition routine, approximately the last 25 seconds of data are compared to the training cases. The result is added to an activity total for both sitting and walking.

6.2.2.2 Feedback Mechanism

Every 30 minutes the activity totals are sent to a server to be logged for the study. These server updates also include a count of notifications that were displayed. Every 60 minutes the totals are checked to gauge the activity level for the last hour. There are two groups that dictate if the user will get a message. User group A receives a message regardless of the user's current activity behavior (one message every hour). Group B only receives a message if the user spent more than 35 minutes sitting. If a message is necessary the user gets an auditory cue prior to the mess age "Ready to get up and move?" This message is accompanied with a stick figure (Figure 6-3) that reflects their current activity level.

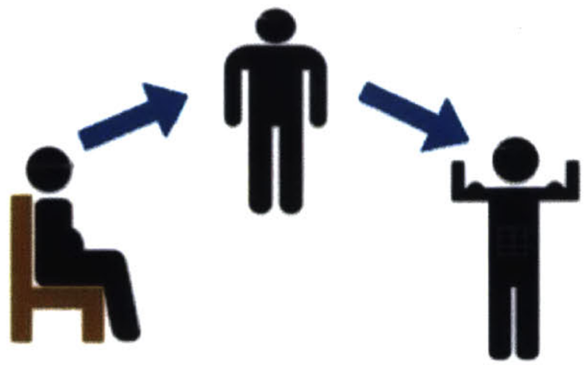


Figure 6-3: Stick figures provided to the user based on their activity level. The arrow dictates the order of progression.

6.2.3 Pilot Study Design

To test the effectiveness of the system for the two groups, a one-week pilot study was conducted with 25 participants. Participants were selected through the use of poster and email solicitations asking for participants wanting to not sit for long periods of time. The majority of participants were graduate students that spent most of the day working on a computer, however a small number of participants were administrative assistants. All participants spent the majority of their workday sitting.

Each participant attended an orientation session in which they completed a pre-survey and a Google Glass orientation. After submitting the pre-survey they were randomly placed into either group A or group B, with A serving as the control group. Each participant was asked to wear the Google Glass with the Move Your Glass application running for at least 3 hours per day. They were also instructed to have their 3 hours occur when they would be doing routine office work.

At the conclusion of the pilot study, each participant returned to complete a post-survey. The post survey asked how many times per day that they received a notification and how often they took a break when the notification appeared, this was also logged with the system. They were also asked why they didn't take a break.

6.2.4 Results

Before conducting a user pilot study, the system was tested for accuracy.

6.2.4.1 Testing Accuracy of System

Testing consisted of two different tests. The first consisted of a user maintaining the same state for an extended period of time. During these tests, Move Your Glass determined the state with 100% accuracy for standing/sitting and running states. During walking states, the system was confused as summarized in Table 6.2. The colors indicate a heat index for the data, illustrating strong relationships in green and weak relationships in red. Yellow reflects a moderate relationship between actual and predicted classes.

		Predicted Class		
		Sitting/Standing	Walking	Running
Actual Class	Walking	55	72	34

Table 6.2: Walking Confusion Matrix. A measure of what state was predicted by the system when a user was in a walking state.

The second test consisted of the user fluently switching between states. Each state was held for a set period of time before transitioning to the next state. The confusion between walking and others states persisted during these tests, however accuracy remained high when determining sedentary and running states. Table 6.3 summarizes the results of this test. Under ideal testing conditions, the Move Your Glass application is able to successfully differentiate between sitting or standing on the one hand, and walking on the other hand using the accelerometer data, therefore these are the two states used in the pilot study. One sample output of this routine is seen in Figure 6-4.

		Predicted Class		
		Sitting/Standing	Walking	Running
Actual Class	Standing	111	0	0
	Walking	8	57	55
	Running	0	0	120

Table 6.3: Mixed states confusion matrix when switching between states

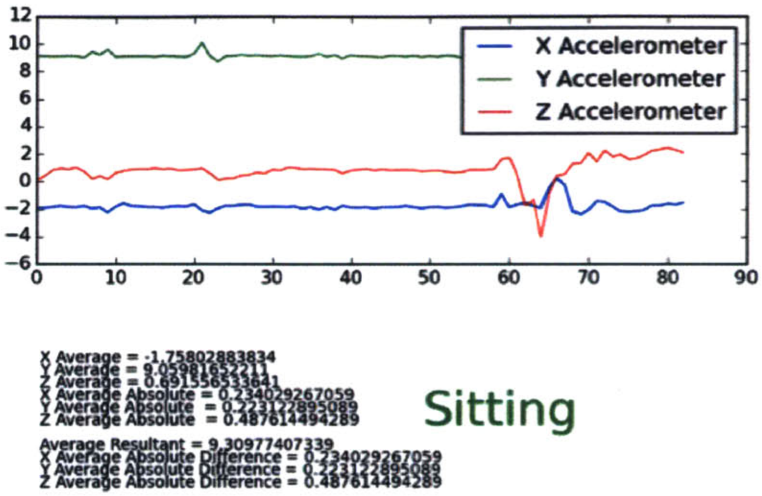


Figure 6-4: Sample activity recognition output for accelerometer data acquired from a sitting subject.

6.2.4.2 User Testing

The time spent sitting each day was measured for each group. Overall, there were 107 data points (70 control, 37 non-control). While on average those in the control group spent less time sitting, this value was not significant (p=.3). A summary of these results can be found in Figure 6-5 where the average time is measured in minutes. Participants in the control group spent less time on average sitting when using the Move Your Glass system meaning that our experiment was not successful. After conducting individual interviews during the post survey, it becomes clear why timing did not have a strong effect.

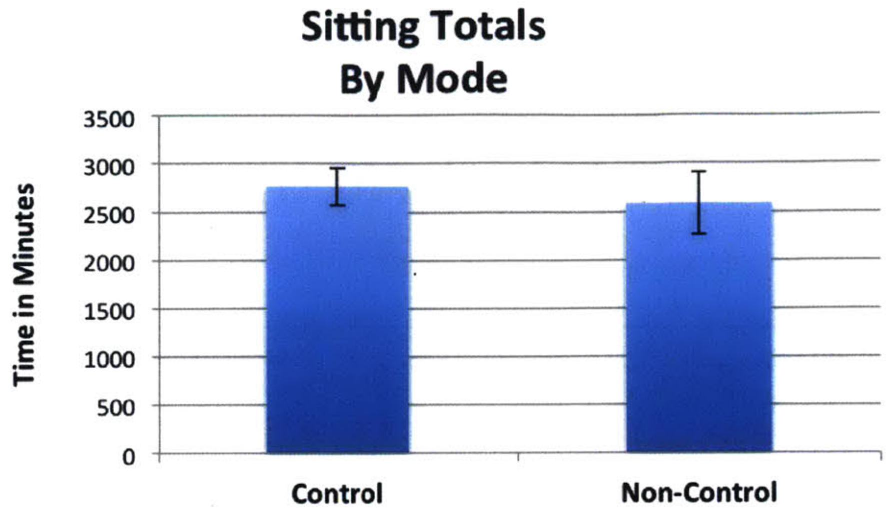


Figure 6-5: Average time spent sitting for both groups

Post experiment surveys were evaluated to gain a better understanding of why the system was not effective. Many users commented on the messages being a good reminder to take a break, however the timing for them was often not opportune. While the messages are triggered after prolonged periods of sitting, their timing is not optimized for the current working state of the user. Many users explained that they would receive a message when they where in the middle of deep thought and didn't want to risk losing their train of thought. A possibly more interesting follow-up experiment would be to time the messages for natural breaks in the user's work.

6.2.5 Discussion

Communicating with Google Glass, an extremely novel device, presented several unique challenges and required a substantial amount of work before any data could be logged from the device for analysis. There are several areas where the system can be improved so that it can be successfully used to induce behavior change.

6.2.5.1 Machine learning

The format of the training data and learning algorithm used for activity recognition could both be further refined. Training data currently has only been collected for one subject and three behaviors. Broadening the pool of activities and test subjects will better train the classification routine. Neural networks and multiple instance-based learning both had high success rates when applied to smartphone accelerometer data, and the algorithms are expected to perform similarly well on Glass. We would also like to investigate the use of other algorithms and methods.

Further work will be done to collect training data from several different subjects as well as larger periods of training data for each subject. Data collection will also include periods of heavy head movements. More data will to improve the robustness and accuracy of Move Your Glass.

6.2.5.2 User Interface

Further development must go into analyzing long-term statistics based on the activities identified and conveying this information back to the Google Glass user on their device. Potential also exists to develop a partnering website to monitor overall data. This could provide the user with summaries of their activities over time and as well as suggestions and tips for improvement.

6.2.5.3 User Engagement

While we recruited users that were motivated to not sit for prolonged periods of time, our system did not provide cues at the best time. Future work could be done to deliver the messages at an opportune moment, for example by taking into account current concentration levels or by monitoring the user's activities and detecting transition points.

6.3 Conclusion

Google Glass is a system that is capable of both activity and behavior monitoring. Systems such as as Move Your Glass that have the ability to give the user just-in-time suggestions are possibly more effective with behavior change than current applications used on other wearables or cellular devices that rely on periodic feedback. Within the pilot study messages given with the Move Your Glass system during moments of high sedentary activity were slightly more effective, although not significantly, than messages given at regular intervals, as seen in Figure 6-5. However, even though users desire to not sit for long periods, productivity is a more effective motivator. Regardless of the timing of the messages, users prioritized their current task over their desire to not sit for prolonged periods. Future systems should target lulls in productivity rather than extended sedentary time as a means of determining optimal timing of message delivery.

Chapter 7

Conclusion

This thesis starts to explore the potential of behavior change triggers delivered on wearable and portable devices. It discussed five different persuasive systems designed for behavior change. Specifically they all offer a real-time trigger for a user to nudge him/her to engage in a healthy behavior. The five systems addressed different aspects of well-being, such as diet, exercise and positive social interaction. The systems were implemented and evaluated in pilot user studies. The studies shed light on the potential of these systems and also revealed some of the difficulties in their effective application. Several lessons can be learned from the design and testing of these systems for future systems. The thesis also describes design considerations when using just-in-time messages for behavior change. When presenting such just-in-time triggers it is important to determine how to get the user's attention and take into account the timing of the message. Finally, the messaging used in the trigger experimented with what content types can be more or less effective.

7.1 Lessons Learned

7.1.1 Fully Utilize Triggers

Triggers are a necessary aspect of persuasive systems, however they can be used as more than simply a reminder for a user to perform an action. If properly used, triggers can motivate users to perform behaviors that they have not already committed to. Triggers should make the desired behavior simple for the user. Messages like "remember to eat healthy" are not as effective as a message suggesting actions the user can take at that moment such as "Hungry? There are some high protein snacks in the next vending machine". By using social networks and competitions triggers can increase motivation or earn buy-in for a new goal. However, competitions inherently have an end where the winner is decided after which users typically return to their old behaviors. Social influence can also only be relied on if at least one member of a social circle is motivated. Motivating members can be challenging, especially if you are trying to get buy-in for an action that is not easy for a user, like interacting with strangers.

7.1.2 Don't Assume Attention

Almost every technology device or application attempts to attract the user's attention. In addition, users often have many competing demands on their attention, especially in a mobile or social context. As a result, users often miss or ignore digital messages. Systems cannot assume that a message presented to a user will be seen and comprehended. This is especially true for new wearable devices. There will always be environments and occasions where a user is so distracted or cognitively depleted that they miss a message. Recognizing these environments and situations is key so that either messages are not displayed during these moments or techniques are developed to overcome these barriers.

7.1.3 Timing Is Key

One of the best strategies for ensuring that a trigger has the user's attention is by utilizing good timing. Triggers should be provided at or just before the moment of decision. They also should not be overused or inaccurate. Inaccuracy refers to messages that are displayed at the wrong time or present inaccurate measurements. Users expect high accuracy from technology and are quick to abandon technology that provides inaccurate information or pesters more than it helps. To ensure adoption it is better to detect behaviors or rely on cues that can be measured with high accuracy. However, timing alone is not enough to overcome users' desire to be productive.

7.1.4 Productivity Is a Strong Motivator

Productivity is a very strong motivator for users. This is evident from the increase in sedentary behavior despite most people being aware of the consequences. Therefore, although users may have goals to change a behavior, in the moment they will chose the action that results in increased productivity rather than health. Systems that simply rely on triggers to remind or motivate a user to perform an action may not be effective for the typical user. Other tactics will need to be utilized such as small changes over a period of time or utilizing a highly motivated social network that the user belongs too. One under explored strategy may be helpful would be to point out the user that increased health and fitness improves overall productivity.

7.1.5 Running Behavior Change User Studies

One of the biggest lessons learned through this thesis is how challenging it is to run well-designed user studies. Generally, participants attempt to meet all of the requirements of the study. However, new technology or lengthy procedures can make it hard for a user to comply. Therefore, it is necessary to design around the challenges that conducting user studies introduce. The following are a summary of lessons learned in regards to designing an effective user study and working with user study participants.

7.1.5.1 Limiting Variables

It is very tempting to design one study to test several parameters at once. However, this typically leads to poorly designed or confusing studies. If positive results are gathered it is often hard to determine what was the cause. It can be easier to design several small studies then one large study that attempts to be all-inclusive. Each study should test one specific and clearly defined hypothesis. Even when that is the case, we learned that when testing systems in real life there are often many unforeseen confounding factors that make it hard to collect meaningful data.

7.1.5.2 Ensure Users Fully Understand the Study Requirements

When using new devices in a user study it is vital that participants know how to properly use a device before starting the study. Ensuring that each participant runs through the steps necessary for the study in a pre-study session. Each participant should also leave with detailed instructions that include screenshots that they can review if they forget a step.

7.1.5.3 Include Several Quality Checks

Even if a participant knows how to follow the steps for a study, that doesn't mean they will fully participate. There should be several ways to verify that each participant is completing all of the required steps. There should also be a protocol for when a participant does not comply. Do they get a warning message or are they excluded from the study? This can help ensure that the study produces clean data.

7.1.5.4 Include an Adjustment Period

The first few days of a study, participants typically struggle to meet all the requirements in the study. Having a built in period that is used to let users get adjusted to the study can ensure that later data are more accurate. Participants should be informed of this adjustment period, as they should try to complete all of the requirements of the study each day and form a routine.

7.1.5.5 Gather Feedback Written and Verbal Post-Survey's

It is common for user studies to end with a survey to assess not only if the studies hypothesis was accurate but also to find shortcomings in the system being tested. However, it is always very helpful to include questions that a proctor asks verbally as part of the post assessment. Depending on the personality of the participant, they may be more forthcoming with information verbally. Where other participants are more willing to give more detailed constructive feedback on paper. Verbal questions also allow for follow-up questions when a users point isn't articulated thoroughly. Regardless, it is typically helpful to quickly review a participant?s response before they leave an clear up any confusing comments.

7.1.5.6 Design for Iterative User Studies

For pretty much all of our experiments we would have benefited from iterating on the study. The first study often reveals unexpected problems and limitations, which necessitate modifications of the system and protocol in a follow-up study.

7.2 Future Work

7.2.1 System Improvements

The systems developed and described in this thesis were built so as to test specific features of triggers such as their timing, content, etc, in relatively short-rem studies. More work is needed to could make these systems effective for long-term behavior change.

SmileCatcher should be made easier to use. This could be done by making an app that takes advantage of the Narrative Clip API that works with the current SmileCatcher website. This would remove the need for users to upload their images and allow the system to provide just-in-time feedback. The Challenge system can be improved by removing the anonymity of uses in the system and by introducing a rolling points system so as to keep the extrinsic motivation high. Users could continue to collect points and cash them in at any point for a prize. This removes the need for the competition to end. Additionally, the makeup of the teams could be varied to explore how mixing social circles can influence game play. Now that we know that productivity is such a strong motivator, Move Your Glass should include productivity measurement to determine better timing of messages. This could include waiting until a user switches tasks or applications. Watch could utilize messaging to help guide users toward effective time management. The system could also aid in setting proper time management goals. These goals could vary over time to more closely match the users' ultimate schedule. Finally, Food Attack needs to be adapted for everyday use cases. This would require food detection or imminent food decision detection to be built in. While there is growing research in this area, detecting food or the imminence of a food decision is still a very hard problem.

7.2.2 Further Research Areas

While this thesis describes several systems that showed promise in persuading users to change a behavior, there are several techniques that could be investigated further.

7.2.2.1 Levels of Motivation

Several of the systems in this thesis were not significantly effective because they battled against user's desire to be productive. With the Challenge system, social circles were identified as an influence that was potentially more impactful than productivity. Further research could investigate other techniques to strengthen a users' motivation to improve their well-being.

7.2.3 Grow with the User

The experiments presented in this thesis only tested short-term behavior change. However in order to maintain that behavior change, systems may have to adapt with the user. Once a habit begins to form systems might need to transition from an encouraging role to a supporting role. This change can take the form of challenging the user with new goals, adding variety to the feedback messaging and adapt to the current state of the user (acquiring a new habit, maintaining a new healthy behavior, working to lose a habit, etc.). One of the problems with current physical activity technologies is their failure to accommodate individual differences and growth [100].

This is especially true for users that continue to find value in wearables after long time use [38]. Long-term value requires long term support.

7.2.3.1 Series of Competitions

This thesis showed that competitions and social networks are effective motivation techniques, however they have their weaknesses. Competitions have an expiration time that makes it likely that users return to previous behaviors. New systems could explore the use of a series of competitions. Similar approaches have been taken with games for adults [101], but they usually require the user to pick and chose their "missions" or challenges. It would be easier to develop a system that could walk a group or individual through a series of competitions to improve a behavior change by applying effective aspects of game theory.

7.2.4 Balancing Social Networks

Social circles have a large impact on social self-esteem and well-being [102]. As an example, the risk of obesity of an individual who has an obese friend or is living with an obese spouse is higher [89]. These effects can work both ways depending on whether the user's social circle is positive or negative. Future work should be done to determine how to effectively create mixed social networks; so unmotivated users can receive the effects of positive influences. Similarly, studies could determine what the maximum number of detrimental members could be in a social network before the group is negatively affected and which members of a social network are most influential. Also a user is a member of a positive reinforcing group at work, but has a negative reinforcing group at home, which will have the largest effect on the user?

Appendix A

Food Attack Appendix

A.1 Food Attack Questionnaire

All questions were presented with a 7 point likert scale, except the question regarding the participants last meal.

1. We hope you enjoyed the snacks, how do you feel about your snack consumption?
2. Do you feel the game influenced your snack consumption? If so, how?
3. When was your last meal?
4. How stressful was your day?
5. How go do you think your reaction time is?

H

Emotional	Fit (Women)	Fit (Men)	Gamification
"Vegetables can reduce your risk of heart disease"	"Vegetables can help with weight loss"	"Vegetables can help with building a lean body	"Vegetables can increase your score"
"Eating fruit may reduce your risk for stroke"	"Eating fruits may help maintain a healthy body weight"	"Eating fruits help maintain optimum health"	"Eating fruits may help improve your score"
"People who eat more protein have more lean muscle"	"People who eat protein eat fewer calories"	"People who eat more protein increases muscle growth"	"People who chose high protein foods score higher"
"Sugar can lead to diabetes"	"Sugar can lead to depression"	"Sugar can lead to premature aging"	"Sugary items can lead to lower scores"
"Eating healthy foods makes you feel good"	"Eating healthy foods makes your waistline smaller"	"Eating healthy foods makes your belly smaller"	"Eating healthy foods make higher scores"

Table A.1: Complete list of messages used in the Food Attack user study

M&M's Remaining	Carrots Remaining	Grapes Remaining	Skittles Remaining	Food	Score	Messages	MessageType	SnackConsumption	Influence	LastMeal	Stress	ReactionTimeRating	Date	Time
9	3	0	9	0	0	0	None	6	Yes	13	2	5	August 10th	11:00 AM
10	10	10	10	0	0	0	None	4	No	3	6	5	August 10th	11:00 AM
7	10	9	10	0	0	0	None	4	No	0	6.6	4.5	August 10th	11:00 AM
8	7	0	8	0	0	0	None	4	No		7	3	August 10th	11:00 AM
0	10	10	0	0	0	0	None	7	Yes	3	7	6	August 10th	11:00 AM
10	10	10	10	0	0	0	None	4	No	0	7	5	August 10th	2:00 PM
10	10	10	10	0	0	0	None	2	No	2	2	6	August 10th	2:00 PM
0	10	0	0	0	0	0	None	4	No	16	6	3	August 10th	2:00 PM
10	10	10	10	0	0	0	None	3	No	0.1	7	4	August 10th	2:00 PM
7	9	0	0	0	0	0	None	5	Yes	2	4	2	August 10th	3:00 PM
1	10	0	2	0	0	0	None	7	Yes	3	7	2	August 10th	3:00 PM
0	0	0	0	0	0	0	None	1	Yes	5	3	2	August 10th	3:00 PM
0	5	0	0	0	0	0	None	7	No	2.5	7	4	August 13th	11:00 AM
0	10	10	0	0	0	0	None	6	Yes	2.5	3	5	August 13th	11:00 AM
2	9	0	9	1	0	0	None	2	No	4	4	2	August 13th	11:00 AM
0	0	0	0	0	0	0	None	6	No	2	6	5.4	August 13th	1:00 PM
10	9	0	9	1	0	0	None	6	No	1	3	3	August 13th	1:00 PM
0	10	0	1	1	0	0	None	7	No	0.2	7	6	August 13th	1:00 PM
0	9	0	0	1	0	0	None	6	Yes	1.5	4	5	August 13th	1:00 PM
6	10	0	0	1	0	0	None	7	No	17	6	5	August 13th	1:00 PM
0	10	3	10	1	0	0	None	6	No	2	6	5	August 13th	2:00 PM
10	3	0	10	1	0	0	None	6	Yes	18	6	5	August 13th	2:00 PM
0	0	0	0	1	0	0	None	6	No	18	5	5	August 13th	2:00 PM
1	5	0	0	1	0	0	None	5	Yes	3	5	4	August 13th	3:30 PM
2	10	10	10	1	0	0	None	3.5	Yes	2	3.5	4.5	August 13th	3:30 PM
10	10	10	10	1	0	0	None	6	No	1	6	3	August 13th	3:30 PM
10	10	10	10	1	0	0	None	5	No	1	7	5	August 14th	11:00 AM
9	6	0	9	1	0	0	None	7	Yes	1	4	2	August 14th	11:00 AM
10	10	10	10	1	0	0	None	4	No	2	4	2	August 14th	11:00 AM
0	0	0	0	1	1	0	None	4	Yes	1	2	3	August 14th	1:00 PM
0	0	0	0	1	1	0	None	6	No	2	5	3	August 14th	1:00 PM
8	10	6	8	1	1	0	None	4	No	3	6	3	August 14th	1:00 PM
6	0	0	8	1	1	0	None	6	No	1	4	4	August 14th	1:00 PM
4	8	0	8	1	1	0	None	7	Yes	1	5	5	August 14th	1:00 PM
10	10	10	10	1	1	0	None	7	No	2.5	6	1	August 14th	2:00 PM
5	1	0	0	1	1	0	None	6	No	4	5	5.5	August 14th	2:00 PM
0	0	0	7	1	1	0	None	7	No	0	4	5	August 14th	2:00 PM
10	10	1	0	1	1	0	None	4	No	1	2	1	August 14th	2:00 PM
8	3	1	4	1	1	0	None	4	No	2	6	5	August 14th	3:00 PM
10	2	1	10	1	1	0	None	6	No	1	5	6	August 14th	3:00 PM
2	1	5	5	1	1	0	None	5.4	Yes	2	2.4	2.3	August 14th	3:00 PM
10	10	10	10	1	1	0	None	4	No	2	4	5	August 14th	3:00 PM
6	6	4	0	1	1	1	E	6	No	2	4	4	August 17th	11:00 AM
10	8	7	10	1	1	1	E	7	Yes	1	6	5.5	August 17th	11:00 AM
8	7	0	10	1	1	1	E	6	Yes	1	6	2	August 17th	11:00 AM
0	5	4	0	1	1	1	E	4	No	0.5	6	1	August 17th	11:00 AM
0	10	8	0	1	1	1	E	7	No	1	2	1	August 17th	2:00 PM
10	10	0	10	1	1	1	E	5	Yes		5	2	August 17th	2:00 PM
0	9	10	0	1	1	1	E	7	No	24	7	2	August 17th	2:00 PM
0	0	0	0	1	1	1	E	4	No	3	7	2	August 17th	3:00 PM
0	0	0	0	1	1	1	G	1	No	0	7	1	Tuesday, August 18th	11:00 AM
10	10	10	10	1	1	1	G	4	Yes	3	3	3	Tuesday, August 18th	11:00 AM
10	0	0	10	1	1	1	G	7	No	18	7	4	Tuesday August 18th	2:00 PM
0	0	0	0	1	1	1	G	5	No	2	7	4	Tuesday August 18th	2:00 PM
0	0	0	0	1	1	1	G	7	Yes	0.5	7	1	Tuesday August 18th	2:00 PM
0	0	0	0	1	1	1	G	4	No	2	3	3	Tuesday August 18th	3:00 PM
0	0	0	0	1	1	1	G	5	No	3.5	6	2	Tuesday August 18th	3:00 PM
10	3	0	10	1	1	1	G	7	Yes	5.5	3.8	2.6	Tuesday August 18th	3:00 PM
0	9	6	5	1	1	1	FM	5	No	16	6	6	Wednesday, August 19th	11:00 AM
0	2	0	0	1	1	1	FW	4	No	5	7	2	Wednesday, August 19th	11:00 AM
0	0	0	0	1	1	1	FM	6	No	1.5	6	4	Wednesday, August 19th	2:00 PM
0	0	0	0	1	1	1	FW	4	No	2.5	6	4	Wednesday, August 19th	2:00 PM
10	9	10	6	1	1	1	FW	5	No	2	3	1	Wednesday, August 19th	2:00 PM
10	10	10	10	1	1	1	FW	3	Yes	2.5	6	4	Wednesday, August 19th	3:00 PM
1	10	0	10	1	1	1	FW	7	No	3	7	3	Wednesday, August 19th	3:00 PM
0	10	0	4	1	1	1	FW	5	No	2	5	5	Wednesday, August 19th	3:00 PM
0	3	0	0	1	1	1	E	4	No	2	5	3	Thursday, October 15th	12:00 PM
7	0	0	4	1	1	1	E	5.8	No	2	5	3	Thursday, October 15th	12:00 PM
10	3	4	10	1	1	1	E	7	Yes	2.5	1	5	Thursday, October 15th	12:00 PM
9	10	1	8	1	1	1	E	6.4	Yes	1	5.5	3.5	Thursday, October 15th	12:00 PM
8	10	0	4	1	1	1	E	5.5	No	17	5.5	1.5	Thursday, October 15th	1:00 PM
3	10	10	0	1	1	1	E	5	Yes	1	7	1	Thursday, October 15th	1:00 PM
10	10	5	7	1	1	1	E	5	No	4	6.2	2.5	Thursday, October 15th	1:00 PM
9	0	0	3	1	1	1	E	7	No	2	5	3	Thursday, October 15th	1:00 PM
10	10	7	10	1	1	1	G	5	No	1.5	2	2	Friday, October 16th	3:30 PM
8	8	10	10	1	1	1	G	3	Yes	3	3	2	Friday, October 16th	3:30 PM
5	0	8	9	1	1	1	G	4	Yes	5.5	5	6	Friday, October 16th	3:30 PM
0	10	3	0	1	1	1	G	7	No	3	1	5	Friday, October 16th	4:30 PM
10	1	9	8	1	1	1	G	7	No	0	2	4	Friday, October 16th	4:30 PM
0	2	5	0	1	1	1	G	6	Yes	4.5	5	6	Friday, October 16th	4:30 PM
5	9	4	4	1	1	1	G	7	No	0.25	7	4	Friday, October 16th	4:30 PM
10	9	8	3	1	1	1	FW	6	Yes		7	5	Wednesday, October 21st	11:00 AM
10	10	0	10	1	1	1	FM	5	No	2	6	4	Wednesday, October 21st	11:00 AM
8	7	7	8	1	1	1	FM	5	Yes	0	4	3	Wednesday, October 21st	12:00 PM
5	8	7	5	1	1	1	FM	6	No	15	4	2	Wednesday, October 21st	12:00 PM
10	10	10	10	1	1	1	FM	7	Yes	16	4	1	Wednesday, October 21st	12:00 PM
0	0	0	0	1	1	1	FW	7	Yes	16	3	3.5	Wednesday, October 21st	12:00 PM
9	10	0	9	1	1	1	FW	6.5	Yes	1	4.5	5.5	Wednesday, October 21st	1:00 PM
6	9	4	3	1	1	1	FM	5	Yes	2	3	5	Wednesday, October 21st	1:00 PM
3	9	0	8	1	1	1	FM	6	No	1	5	2	Wednesday, October 21st	1:00 PM

Table A.2: Food Attack Raw Data

Appendix B

SmileCatcher Appendix

B.1 SmileCatcher Pre-Study Questions

1. How many people do you generally see in a day (estimate is fine)?
2. How many people do you interact with on a daily basis (estimate) (circle one)?
3. Which of the following most effectively effect your mood for the day
   (a) Work
   (b) Family or Friends
   (c) Mood of others
   (d) Day of the week (weekday vs weekend or Monday vs. Friday)
   (e) Workouts
   (f) Amount of sleep
4. How many people in your daily life do you attempt to make happy?
5. What categories do these people fall into (check all that apply)?
   (a) Family member
   (b) Spouse
   (c) Friend

(d) Co-worker

(e) Stranger

(f) Other (please specify)

6. Do you ever reflect on the people you made happy during a given day (circle one)?

B.2 SmileCatcher Post-Study Questions

1. How difficult was it to meet your smile goal?

2. On average how many smiles did you collect each day?

3. Do you feel the number of smiles you collected each day had any effect on your mood? Why or why not?

4. When collecting smiles which categories of people did you target?

   (a) Family member

   (b) Spouse

   (c) Friend

   (d) Co-worker

   (e) Stranger

   (f) Other (please specify)

5. Did you review the smiles you collected? Why or why not?

Table B.1: SmileCatcher Raw Data

User number	ID Code	Session	User	First Week Smiles	Second Week Smiles	Increase in Smiles?	Number of images uploaded
12	79207	2	1	18	24	TRUE	12459
11	68095	2	2	13	18	TRUE	4314
21	66658	3	3	3	11	TRUE	3855
22	78370	3	4	19	4	FALSE	10454
114	42545	1	5	14	13	FALSE	3431
116	34509	1	6	14	33	TRUE	1574
113	13464	1	7	8	13	TRUE	2146
112	34245	1	8	2	41	TRUE	1964
117	34235	1	9	0	6	TRUE	2195
108	34568	1	10	7	3	FALSE	1693
110	49796	1	11	8	13	TRUE	12024
107	14789	1	12	13	11	FALSE	4235
109	35683	1	13	0	16	TRUE	3567
26	81550	3	14	1	4	TRUE	4326
14	63334	2	15	1	0	FALSE	2824
15	32392	2	16	0	0	FALSE	2585
23	82450	3	17	0	1	TRUE	2089
18	76366	3	18	3	1	FALSE	1795
19	68161	3	19	1	2	TRUE	1470
111	12442	1	20	0	3	TRUE	3246
10	72166	2	21	0	0	FALSE	7
13	35167	2	22	0	0	FALSE	18
16	78367	2	23	0	0	FALSE	40
8	68350	2	24	0	0	FALSE	0
20	77566	3	25	0	1	TRUE	90
24	51862	3	26	0	0	FASE	0
17	82807	3	27	0	0	FALSE	18

Appendix C

Attention Grabber Appendix

Table C.1: Attention Grabber Study Part 1 Raw Data

Person	Action 1 No	Action 1 sound	Action 1 visual	Action 2 No	Action 2 sound	Action 2 visual	Action 3 No	Action 3 sound	Action 3 visual
1	1	1	1	0	0	1	1	1	1
2	0	1	0	0	1	1	1	1	1
3	1	1	1	0	0	0	1	1	1
4	1	0	1	0	1	0	1	1	1
5	1	0	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	0	0	1	1
8	1	1	1	1	1	1	1	1	1
9	0	0	0	1	1	0	1	1	1
10	1	1	1	0	1	1	1	1	1

Table C.2: Attention Grabber Study Part 2 Raw Data

id	time	intervention	remember_1	remember_2	remember_3	remember_4	remember_5	look_all_the_time	count_think_saw
1	1:21 PM	audio	1	0	0	1	1	0	5
2	1:25 PM	blink	1	0	1	0	1	0	4
3	2:00 PM	audio	1	1	0	0	1	0	3
4	2:25 PM	audio	1	1	1	1	1	0	5
5	2:39 PM	blink	0	0	0	0	0	0	0
6	2:49 PM	audio	1	1	0	1	1	0	5
7	2:52 PM	blink	0	0	0	0	0	0	0
8	4:17 PM	blink	0	0	0	0	0	0	4.5
9	4:38 PM	audio	0	0	1	0	1	0	4.5
10	4:50 PM	blink	0	1	0	0	1	0	5
11	5:17 PM	audio	1	1	0	0	1	0	2.5
12	5:28 PM	blink	0	0	0	0	0	0	0
13	5:35 PM	audio	1	1	0	0	0	0	5.5
14	5:53 PM	audio	0	0	0	0	0	0	2.5
15	6:08 PM	blink	0	0	0	0	0	0	0
16	6:14 PM	blink	1	0	0	1	0	0	4
17	6:34 PM	audio	1	1	0	0	1	1	6
18	6:43 PM	blink	1	1	0	0	1	0	5
19	6:58 PM	audio	0	0	0	0	0	0	0
20	7:03 PM	blink	0	0	1	0	0	0	5.5
21	7:08 PM	audio	0	1	0	0	0	0	4.5
22	7:14 PM	blink	1	0	0	0	0	1	3
23	7:17 PM	audio	0	0	0	0	0	0	0
24	7:24 PM	blink	0	0	0	0	0	0	0
25	7:30 PM	audio	0	0	0	0	0	0	0
26	7:35 PM	blink	1	0	1	0	0	0	0
27	7:51 PM	audio	0	0	0	0	0	0	0
28	7:58 PM	blink	0	0	0	0	0	0	0
29	8:10 PM	audio	1	1	1	0	0	0	5
30	8:20 PM	blink	1	0	0	0	0	0	5

Appendix D

Move Your Glass Appendix

D.1 Move Your Glass Pre-Study Questions

1. What percentage of your day (excluding time spent sleeping) would you like to spend sitting?

    (a) 0-10%

    (b) 10-30%

    (c) 30-50%

    (d) 50-70%

    (e) Over 70%

2. What percentage of your day (excluding time spent sleeping) would you like to spend walking?

    (a) 0-10%

    (b) 10-30%

    (c) 30-50%

    (d) 50-70%

    (e) Over 70%

3. What percentage of your day (excluding time spent sleeping) would you like to spend running?
    (a) 0-10%
    (b) 10-30%
    (c) 30-50%
    (d) 50-70%
    (e) Over 70%
4. What age group do you belong in?
    (a) 18-25
    (b) 25-40
    (c) 40-65
    (d) 65 and over
5. What is your gender?
6. How often do you workout each week?
    (a) I don't
    (b) 1-2 times per week
    (c) 3-4 times per week
    (d) More than 4 times per week

D.2 Move Your Glass Post-Survey Questions

1. About how many times a day did you receive a message to stand up?
    (a) 0
    (b) 1-2

(c) 3-4

(d) More than 4

2. How long each day did you wear the Google Glass?

   (a) 0-2 hours

   (b) 2-4 hours

   (c) 4-6 hours

   (d) More than 6 hours

3. If you received messages, were they effective at having you take a break?

4. Why or why not?

Table D.1: Move your Glass Raw Data

id	time	intervention	remember_1	remember_2	remember_3	remember_4	remember_5	look_all_the_time	count_think_saw
1	1:21 PM	audio	1	0	0	1	1	0	5
2	1:25 PM	blink	1	0	1	0	1	0	4
3	2:00 PM	audio	1	1	0	0	1	0	3
4	2:25 PM	audio	1	1	1	1	1	0	5
5	2:39 PM	blink	0	0	0	0	0	0	0
6	2:49 PM	audio	1	1	0	1	1	0	5
7	2:52 PM	blink	0	0	0	0	0	0	0
8	4:17 PM	blink	0	0	0	0	0	0	4.5
9	4:38 PM	audio	0	0	1	0	1	0	4.5
10	4:50 PM	blink	0	1	0	0	1	0	5
11	5:17 PM	audio	1	1	0	0	1	0	2.5
12	5:28 PM	blink	0	0	0	0	0	0	0
13	5:35 PM	audio	1	1	0	0	0	0	5.5
14	5:53 PM	audio	0	0	0	0	0	0	2.5
15	6:08 PM	blink	0	0	0	0	0	0	0
16	6:14 PM	blink	1	0	0	1	0	0	4
17	6:34 PM	audio	1	1	0	0	1	1	6
18	6:43 PM	blink	1	1	0	0	1	0	5
19	6:58 PM	audio	0	0	0	0	0	0	0
20	7:03 PM	blink	0	0	1	0	0	0	5.5
21	7:08 PM	audio	0	1	0	0	0	0	4.5
22	7:14 PM	blink	1	0	0	0	0	1	3
23	7:17 PM	audio	0	0	0	0	0	0	0
24	7:24 PM	blink	0	0	0	0	0	0	0
25	7:30 PM	audio	0	0	0	0	0	0	0
26	7:35 PM	blink	1	0	1	0	0	0	0
27	7:51 PM	audio	0	0	0	0	0	0	0
28	7:58 PM	blink	0	0	0	0	0	0	0
29	8:10 PM	audio	1	1	1	0	0	0	5
30	8:20 PM	blink	1	0	0	0	0	0	5

Appendix E

Watch Appendix

E.1 Watch Pre-Survey Questions

1. What time management tools do you use?

2. What are some of the things that you do/ do not like about those tools?

3. On the scale below, mark how true the following statement is for you: I wish that I could improve the way I spend my time.

4. On the scale below, mark how true the following statement is for you: I often feel that I do not have enough time to accomplish everything that I hoped to accomplish in the day.

5. On the scale below, mark how true the following statement is for you: I am consistently able to make time for the things that are important to me.

6. On a typical day, how often do your meet your workout time goals?

7. On a typical day, how often to you get as much sleep as you planned?

8. On a typical day, how often to you meet your goal for spending time at work?

9. On a typical day, how often to you meet your goal for spending time at home?

E.2 Watch Post-Survey Questions

1. How easy was it to use the goal tracker app?
2. What problems (if any) did you encounter with the goal tracker app?
3. Were there any features that the goal tracker app should have included?
4. Did you use any time management tools in addition to the goal tracker this past week? If yes what?
5. How often did you check the goal tracker app? Every ? (e.g. hours, minutes, etc.)
6. Do you feel that using WATCH made it easier to stay on track? Why or why not?
7. Is the goal tracker app?s interface too complex to be effectively used as a time management tool?
8. Is the goal tracker app?s interface too simple to be an effective time management tool?
9. Did using the goal tracker app allow you to make more realistic goals for the day?
10. Did outlining time management goals help you make more realistic goals for your day?
11. On a typical day, how often do your meet your workout time goals?
12. On a typical day, how often to you get as much sleep as you planned?
13. On a typical day, how often to you meet your goal for spending time at work?
14. On a typical day, how often to you meet your goal for spending time at home?

ıail.com	1/27/16	780	0	780	600	0	600	60	1440	1380	420	376	44	600
ıail.com	1/28/16	780	0	780	600	0	600	60	1440	1380	420	379	41	600
ıail.com	1/29/16	780	24	756	600	0	600	60	1416	1356	420	348	72	600
ıail.com	1/30/16	780	0	780	600	0	600	60	1440	1380	420	350	70	600
ıail.com	1/31/16	780	673	107	600	0	600	60	767	707	420	491	71	600
ıail.com	2/1/16	780	0	780	600	575	25	60	865	805	420	305	115	600
ıail.com	2/2/16	780	0	780	600	224	376	60	1216	1156	420	0	420	600
ıail.com	2/3/16	780	0	780	600	0	600	60	1440	1380	420	412	8	600
ıail.com	2/4/16	780	0	780	600	360	240	60	1080	1020	420	540	120	600
ıdo@gmail.com	1/22/16	240	765	525	480	135	345	720	540	180	480	360	120	180
ıdo@gmail.com	1/23/16	240	1153	913	480	210	270	720	77	643	480	428	52	180
ıdo@gmail.com	1/24/16	240	837	597	480	90	390	720	513	207	480	320	160	180
ıdo@gmail.com	1/25/16	240	1035	795	480	45	435	720	360	360	480	325	155	180
ıdo@gmail.com	1/26/16	240	1038	798	480	90	390	720	312	408	480	335	145	180
ıdo@gmail.com	1/27/16	240	1381	1141	480	0	480	720	59	661	480	180	300	180
ıdo@gmail.com	1/28/16	240	1219	979	480	0	480	720	221	499	480	0	480	180
ıdo@gmail.com	1/29/16	240	540	300	480	0	480	720	900	180	480	0	480	180
ıdo@gmail.com	1/30/16	240	989	749	480	0	480	720	451	269	480	0	480	180
ıdo@gmail.com	1/31/16	240	973	733	480	0	480	720	467	253	480	0	480	180
ıdo@gmail.com	2/1/16	240	0	240	480	0	480	720	1440	720	480	310	170	180
ıdo@gmail.com	2/2/16	240	250	10	480	0	480	720	1190	470	480	0	480	180
ıdo@gmail.com	2/3/16	240	858	618	480	0	480	720	582	138	480	0	480	180
ıdo@gmail.com	2/4/16	240	744	504	480	80	400	720	616	104	480	445	35	180
ıdo@gmail.com	2/5/16	240	950	710	480	54	426	720	436	284	480	0	480	180
ıdo@gmail.com	2/6/16	240	674	434	480	0	480	720	766	46	480	0	480	180
ıdo@gmail.com	2/7/16	240	1131	891	480	0	480	720	309	411	480	445	35	180
ıdo@gmail.com	2/8/16	240	757	517	480	55	425	720	628	92	480	340	140	180
@gmail.com	1/22/16	900	0	900	360	403	43	180	1037	857	480	369	111	120
@gmail.com	1/23/16	900	0	900	360	0	360	180	1440	1260	480	425	55	120
@gmail.com	1/24/16	900	0	900	360	0	360	180	1440	1260	480	426	54	120
@gmail.com	1/25/16	900	0	900	360	15	345	180	1425	1245	480	410	70	120
@gmail.com	1/26/16	900	0	900	360	0	360	180	1440	1260	480	297	183	120
@gmail.com	1/27/16	900	293	607	360	0	360	180	1147	967	480	541	61	120
@gmail.com	1/28/16	900	764	136	360	0	360	180	676	496	480	484	4	120
@gmail.com	1/29/16	900	681	219	360	0	360	180	759	579	480	510	30	120
@gmail.com	1/30/16	900	0	900	360	0	360	180	1440	1260	480	506	26	120
@gmail.com	1/31/16	900	1109	209	360	0	360	180	331	151	480	510	30	120
@gmail.com	2/1/16	900	1409	509	360	0	360	180	31	149	480	420	60	120
@gmail.com	2/2/16	900	664	236	360	224	136	180	552	372	480	410	70	120
@gmail.com	2/3/16	900	884	16	360	0	360	180	556	376	480	492	12	120
@gmail.com	2/4/16	900	1245	345	360	0	360	180	195	15	480	333	147	120
@gmail.com	2/5/16	900	718	182	360	0	360	180	722	542	480	0	480	120
@gmail.com	2/6/16	900	195	705	360	0	360	180	1245	1065	480	0	480	120
@gmail.com	2/7/16	900	1214	314	360	0	360	180	226	46	480	0	480	120
@gmail.com	2/8/16	900	899	1	360	29	331	180	512	332	480	0	480	120
@gmail.com	2/9/16	900	928	28	360	89	271	180	423	243	480	0	480	120
@gmail.com	2/10/16	900	733	167	360	0	360	180	707	527	480	0	480	120
@gmail.com	2/11/16	900	0	900	360	0	360	180	1440	1260	480	0	480	120
@gmail.com	2/12/16	900	0	900	360	0	360	180	1440	1260	480	0	480	120
@gmail.com	2/13/16	900	0	900	360	0	360	180	1440	1260	480	0	480	120
@gmail.com	2/14/16	900	117	783	360	0	360	180	1323	1143	480	0	480	120
@gmail.com	2/15/16	900	0	900	360	0	360	180	1440	1260	480	0	480	120
gmail.com	1/22/16	960	0	960	540	164	376	-60	1276	1336	480	436	44	180
gmail.com	1/23/16	960	0	960	540	0	540	-60	1440	1500	480	498	18	180
gmail.com	1/24/16	960	0	960	540	0	540	-60	1440	1500	480	441	39	180
gmail.com	1/25/16	960	0	960	540	456	84	-60	984	1044	480	440	40	180
gmail.com	1/26/16	960	95	865	540	194	346	-60	1151	1211	480	428	52	180
gmail.com	1/27/16	960	968	8	540	194	346	-60	278	338	480	401	79	180
gmail.com	1/28/16	960	807	153	540	0	540	-60	633	693	480	554	74	180
gmail.com	1/29/16	960	430	530	540	0	540	-60	1010	1070	480	470	10	180
gmail.com	1/30/16	960	0	960	540	0	540	-60	1440	1500	480	476	4	180
gmail.com	1/31/16	960	86	874	540	0	540	-60	1354	1414	480	444	36	180
gmail.com	2/1/16	960	674	286	540	465	75	-60	301	361	480	480	0	180
gmail.com	2/2/16	960	824	136	540	434	106	-60	182	242	480	484	4	180

:@gmail.com	1/27/16	0	387	387	480	0	480	960	1053	93	480	433	47	90
:@gmail.com	1/28/16	0	0	0	480	0	480	960	1440	480	480	474	6	90
:@gmail.com	1/29/16	0	0	0	480	0	480	960	1440	480	480	391	89	90
:@gmail.com	1/30/16	0	0	0	480	0	480	960	1440	480	480	359	121	90
:@gmail.com	1/31/16	0	0	0	480	0	480	960	1440	480	480	430	50	90
:@gmail.com	2/1/16	0	0	0	480	0	480	960	1440	480	480	421	59	90
:@gmail.com	2/2/16	0	0	0	480	0	480	960	1440	480	480	371	109	90
:@gmail.com	2/3/16	0	0	0	480	0	480	960	1440	480	480	484	4	90
:@gmail.com	2/4/16	0	0	0	480	0	480	960	1440	480	480	318	162	90
:@gmail.com	2/5/16	0	0	0	480	0	480	960	1440	480	480	406	74	90
:@gmail.com	2/6/16	0	0	0	480	0	480	960	1440	480	480	686	206	90
:@gmail.com	2/7/16	0	0	0	480	0	480	960	1440	480	480	318	162	90
:@gmail.com	2/8/16	0	0	0	480	0	480	960	1440	480	480	398	82	90
:@gmail.com	2/9/16	0	112	112	480	0	480	960	1328	368	480	384	96	90
:@gmail.com	2/10/16	0	610	610	480	0	480	960	830	130	480	428	52	90
:@gmail.com	2/11/16	0	927	927	480	0	480	960	513	447	480	330	150	90
:@gmail.com	2/12/16	0	1137	1137	480	0	480	960	303	657	480	353	127	90
:@gmail.com	2/13/16	0	0	0	480	0	480	960	1440	480	480	0	480	90
:@gmail.com	2/14/16	0	1064	1064	480	0	480	960	376	584	480	435	45	90
:@gmail.com	2/15/16	0	1229	1229	480	0	480	960	211	749	480	458	22	90
:@gmail.com	2/16/16	0	1019	1019	480	0	480	960	421	539	480	435	45	90
:@gmail.com	2/17/16	0	824	824	480	0	480	960	616	344	480	381	99	90
:@gmail.com	2/18/16	0	719	719	480	15	465	960	706	254	480	431	49	90
:@gmail.com	2/19/16	0	914	914	480	0	480	960	526	434	480	417	63	90
;mail.com	1/28/16	300	0	300	480	0	480	660	1440	780	450	428	22	180
;mail.com	1/29/16	300	547	247	480	0	480	660	893	233	450	452	2	180
;mail.com	1/30/16	300	0	300	480	0	480	660	1440	780	450	517	67	180
;mail.com	1/31/16	300	0	300	480	0	480	660	1440	780	450	367	83	180
;mail.com	2/1/16	300	0	300	480	0	480	660	1440	780	450	520	70	180
;mail.com	2/2/16	300	1366	1066	480	0	480	660	74	586	450	448	2	180
;mail.com	2/3/16	300	0	300	480	0	480	660	1440	780	450	395	55	180
;mail.com	2/4/16	300	0	300	480	0	480	660	1440	780	450	497	47	180
;mail.com	2/5/16	300	0	300	480	0	480	660	1440	780	450	405	45	180
;mail.com	2/6/16	300	0	300	480	0	480	660	1440	780	450	362	88	180
;mail.com	2/7/16	300	1217	917	480	0	480	660	223	437	450	521	71	180
;mail.com	2/8/16	300	0	300	480	0	480	660	1440	780	450	402	48	180
;mail.com	2/9/16	300	0	300	480	0	480	660	1440	780	450	394	56	180
;mail.com	2/10/16	300	0	300	480	0	480	660	1440	780	450	494	44	180
;mail.com	2/11/16	300	0	300	480	0	480	660	1440	780	450	425	25	180
;mail.com	1/28/16	660	0	660	180	0	180	600	1440	840	450	491	41	420
;mail.com	1/29/16	660	0	660	180	0	180	600	1440	840	450	148	302	420
;mail.com	1/30/16	660	1060	400	180	0	180	600	380	220	450	875	425	420
;mail.com	1/31/16	660	185	475	180	0	180	600	1255	655	450	125	325	420
;mail.com	2/1/16	660	135	525	180	0	180	600	1305	705	450	447	3	420
;mail.com	2/2/16	660	0	660	180	0	180	600	1440	840	450	535	85	420
;mail.com	2/3/16	660	0	660	180	0	180	600	1440	840	450	570	120	420
;mail.com	2/4/16	660	90	570	180	0	180	600	1350	750	450	40	410	420
;mail.com	2/5/16	660	0	660	180	0	180	600	1440	840	450	0	450	420
;mail.com	2/6/16	660	0	660	180	0	180	600	1440	840	450	164	286	420
;mail.com	2/7/16	660	0	660	180	0	180	600	1440	840	450	648	198	420
;mail.com	2/8/16	660	1104	444	180	0	180	600	336	264	450	453	3	420
;mail.com	2/9/16	660	297	363	180	0	180	600	1143	543	450	0	450	420
;mail.com	2/10/16	660	303	357	180	0	180	600	1137	537	450	54	396	420
;mail.com	2/11/16	660	759	99	180	0	180	600	681	81	450	404	46	420
r@gmail.com	1/28/16	480	0	480	420	74	346	540	1366	826	480	904	424	60
r@gmail.com	1/29/16	480	0	480	420	0	420	540	1440	900	480	364	116	60
r@gmail.com	1/30/16	480	0	480	420	0	420	540	1440	900	480	543	63	60
r@gmail.com	1/31/16	480	184	296	420	0	420	540	1256	716	480	433	47	60
r@gmail.com	2/1/16	480	1019	539	420	0	420	540	421	119	480	518	38	60
r@gmail.com	2/2/16	480	929	449	420	285	135	540	226	314	480	483	3	60
r@gmail.com	2/3/16	480	1169	689	420	120	300	540	151	389	480	519	39	60
r@gmail.com	2/4/16	480	781	301	420	105	315	540	554	14	480	471	9	60
r@gmail.com	2/5/16	480	576	96	420	330	90	540	534	6	480	448	32	60
r@gmail.com	2/6/16	480	540	60	420	0	420	540	900	360	480	535	55	60

mail.com	2/1/16	480	820	340	480	0	480	480	620	140	450	450	0	240
mail.com	2/2/16	480	689	209	480	0	480	480	751	271	450	483	33	240
mail.com	2/3/16	480	692	212	480	0	480	480	748	268	450	467	17	240
mail.com	2/4/16	480	776	296	480	0	480	480	664	184	450	325	125	240
mail.com	2/5/16	480	1349	869	480	0	480	480	91	389	450	684	234	240
mail.com	2/6/16	480	0	480	480	0	480	480	1440	960	450	611	161	240
mail.com	2/7/16	480	1274	794	480	0	480	480	166	314	450	631	181	240
mail.com	2/8/16	480	827	347	480	0	480	480	613	133	450	450	0	240
mail.com	2/9/16	480	1268	788	480	0	480	480	172	308	450	563	113	240
mail.com	2/10/16	480	774	294	480	0	480	480	666	186	450	428	22	240
mail.com	2/11/16	480	1044	564	480	0	480	480	396	84	450	468	18	240
il.com	1/30/16	480	0	480	480	0	480	480	1440	960	480	0	480	240
il.com	1/31/16	480	0	480	480	441	39	480	999	519	480	0	480	240
il.com	2/1/16	480	494	14	480	436	44	480	510	30	480	0	480	240
il.com	2/2/16	480	0	480	480	719	239	480	721	241	480	0	480	240
il.com	2/3/16	480	58	422	480	832	352	480	550	70	480	0	480	240
il.com	2/4/16	480	0	480	480	644	164	480	796	316	480	0	480	240
il.com	2/5/16	480	0	480	480	835	355	480	605	125	480	0	480	240
il.com	2/6/16	480	0	480	480	906	426	480	534	54	480	0	480	240
il.com	2/7/16	480	0	480	480	683	203	480	757	277	480	0	480	240
il.com	2/8/16	480	104	376	480	0	480	480	1336	856	480	0	480	240
il.com	2/9/16	480	109	371	480	0	480	480	1331	851	480	123	357	240
il.com	2/10/16	480	628	148	480	0	480	480	812	332	480	436	44	240
il.com	2/11/16	480	461	19	480	106	374	480	873	393	480	453	27	240
il.com	2/12/16	480	312	168	480	0	480	480	1128	648	480	101	379	240
il.com	2/13/16	480	332	148	480	0	480	480	1108	628	480	317	163	240
il.com	2/14/16	480	0	480	480	352	128	480	1088	608	480	7	473	240
il.com	2/15/16	480	0	480	480	428	52	480	1012	532	480	699	219	240
il.com	2/16/16	480	0	480	480	1209	729	480	231	249	480	540	60	240
il.com	2/17/16	480	0	480	480	686	206	480	754	274	480	490	10	240
il.com	2/18/16	480	0	480	480	1007	527	480	433	47	480	145	335	240
il.com	2/19/16	480	0	480	480	692	212	480	748	268	480	417	63	240
il.com	2/20/16	480	0	480	480	1138	658	480	302	178	480	847	367	240
il.com	2/21/16	480	0	480	480	1122	642	480	318	162	480	1109	629	240
mail.com	1/30/16	180	602	422	480	0	480	780	838	58	420	606	186	120
mail.com	1/31/16	180	0	180	480	0	480	780	1440	660	420	947	527	120
mail.com	2/1/16	180	1214	1034	480	180	300	780	46	734	420	457	37	120
mail.com	2/2/16	180	1064	884	480	344	136	780	32	748	420	470	50	120
mail.com	2/3/16	180	809	629	480	540	60	780	91	689	420	374	46	120
mail.com	2/4/16	180	884	704	480	390	90	780	166	614	420	449	29	120
mail.com	2/5/16	180	659	479	480	585	105	780	196	584	420	486	66	120
mail.com	2/6/16	180	0	180	480	0	480	780	1440	660	420	350	70	120
mail.com	2/7/16	180	1409	1229	480	0	480	780	31	749	420	462	42	120
mail.com	2/8/16	180	0	180	480	0	480	780	1440	660	420	725	305	120
mail.com	2/9/16	180	575	395	480	374	106	780	491	289	420	546	126	120
mail.com	2/10/16	180	295	115	480	0	480	780	1145	365	420	430	10	120
mail.com	2/11/16	180	0	180	480	0	480	780	1440	660	420	471	51	120
mail.com	2/12/16	180	0	180	480	0	480	780	1440	660	420	514	94	120
mail.com	2/13/16	180	1168	988	480	0	480	780	272	508	420	460	40	120
mail.com	2/14/16	180	0	180	480	0	480	780	1440	660	420	90	330	120
mail.com	2/15/16	180	0	180	480	0	480	780	1440	660	420	631	211	120
nail.com	1/30/16	300	0	300	480	0	480	660	1440	780	480	610	130	120

Table E.1: Watch Raw Data

Bibliography

[1] B. J. Fogg, "A behavior model for persuasive design," in *Proceedings of the 4th international Conference on Persuasive Technology.* ACM, 2009, p. 40.

[2] C. Duhigg, *The power of habit: Why we do what we do in life and business.* Random House, 2012, vol. 34, no. 10.

[3] E. K. Choe, N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz, "Understanding quantified-selfers' practices in collecting and exploring personal data," in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems.* ACM, 2014, pp. 1143–1152.

[4] D. Ariely and G. Loewenstein, "The heat of the moment: The effect of sexual arousal on sexual decision making," *Journal of Behavioral Decision Making*, vol. 19, no. 2, pp. 87–98, 2006.

[5] E. L. Deci and R. M. Ryan, "Facilitating optimal motivation and psychological well-being across life's domains." *Canadian Psychology/Psychologie canadienne*, vol. 49, no. 1, p. 14, 2008.

[6] C. Fernet, F. Guay, and C. Senécal, "Adjusting to job demands: The role of work self-determination and job control in predicting burnout," *Journal of vocational behavior*, vol. 65, no. 1, pp. 39–56, 2004.

[7] L. G. Pelletier, S. C. Dion, M. Slovinec-D'Angelo, and R. Reid, "Why do you regulate what you eat? relationships between forms of regulation, eating behaviors, sustained dietary behavior change, and psychological adjustment," *Motivation and Emotion*, vol. 28, no. 3, pp. 245–277, 2004.

[8] E. L. Deci, R. Koestner, and R. M. Ryan, "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation." *Psychological bulletin*, vol. 125, no. 6, p. 627, 1999.

[9] E. L. Deci and R. M. Ryan, *Intrinsic motivation.* Wiley Online Library, 1975.

[10] V. J. Strecher, B. M. DeVellis, M. H. Becker, and I. M. Rosenstock, "The role of self-efficacy in achieving health behavior change," *Health Education & Behavior*, vol. 13, no. 1, pp. 73–92, 1986.

[11] T. Kasser and R. M. Ryan, "Further examining the american dream: Differential correlates of intrinsic and extrinsic goals," *Personality and Social Psychology Bulletin*, vol. 22, no. 3, pp. 280–287, 1996.

[12] K. M. Sheldon, R. M. Ryan, E. L. Deci, and T. Kasser, "The independent effects of goal contents and motives on well-being: It?s both what you pursue and why you pursue it," *Personality and Social Psychology Bulletin*, vol. 30, no. 4, pp. 475–486, 2004.

[13] M. Vansteenkiste, J. Simons, B. Soenens, W. Lens *et al.*, "How to become a persevering exerciser? providing a clear, future intrinsic goal in an autonomy-supportive way," *Journal of Sport and exercise Psychology*, vol. 26, no. 2, pp. 232–249, 2004.

[14] V. J. Strecher, G. H. Seijts, G. J. Kok, G. P. Latham, R. Glasgow, B. DeVellis, R. M. Meertens, and D. W. Bulger, "Goal setting as a strategy for health behavior change," *Health Education & Behavior*, vol. 22, no. 2, pp. 190–200, 1995.

[15] B. J. Fogg, "Persuasive technology: using computers to change what we think and do," *Ubiquity*, vol. 2002, no. December, p. 5, 2002.

[16] R. Bridgelall, "Hybrid bluetooth/rfid based real time location tracking," Apr. 6 2004, uS Patent 6,717,516.

[17] M. A. Case, H. A. Burwick, K. G. Volpp, and M. S. Patel, "Accuracy of smartphone applications and wearable devices for tracking physical activity data," *Jama*, vol. 313, no. 6, pp. 625–626, 2015.

[18] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 6, no. 2, pp. 156–166, 2005.

[19] J. Whitehill, G. Littlewort, I. Fasel, M. Bartlett, and J. Movellan, "Toward practical smile detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 11, pp. 2106–2111, 2009.

[20] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of emotion recognition using facial expressions, speech and multimodal information," in *Proceedings of the 6th international conference on Multimodal interfaces.* ACM, 2004, pp. 205–211.

[21] M. Wunsch, A. Stibe, A. Millonig, S. Seer, C. Dai, K. Schechtner, and R. C. Chin, "What makes you bike? exploring persuasive strategies to encourage low-energy mobility," in *Persuasive Technology.* Springer, 2015, pp. 53–64.

[22] M. Kaptein, B. De Ruyter, P. Markopoulos, and E. Aarts, "Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 2, no. 2, p. 10, 2012.

[23] S. B. Gram-Hansen and T. Ryberg, "Acttention–influencing communities of practice with persuasive learning designs," in *Persuasive Technology.* Springer, 2015, pp. 184–195.

[24] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive computing.* Springer, 2004, pp. 1–17.

[25] E. M. Tapia, S. S. Intille, and K. Larson, *Activity recognition in the home using simple and ubiquitous sensors.* Springer, 2004.

[26] S. A. Munson and S. Consolvo, "Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on.* IEEE, 2012, pp. 25–32.

[27] S. Consolvo, P. Klasnja, D. W. McDonald, and J. A. Landay, "Goal-setting considerations for persuasive technologies that encourage physical activity," in *Proceedings of the 4th international Conference on Persuasive Technology.* ACM, 2009, p. 8.

[28] M. Maheshwari, S. Chatterjee, and D. Drew, "Exploring the persuasiveness of ?just-in-time? motivational messages for obesity management," in *Persuasive Technology.* Springer, 2008, pp. 258–261.

[29] A. Kadomura, K. Tsukada, and I. Siio, "Educatableware: computer-augmented tableware to enhance the eating experiences," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems.* ACM, 2013, pp. 3071–3074.

[30] Darma. (2015) Sit smart for a healthy body and mind. [Online]. Available: http://darma.co

[31] E. Arroyo, L. Bonanni, and T. Selker, "Waterbot: exploring feedback and persuasive techniques at the sink," in *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 2005, pp. 631–639.

[32] J.-C. Ko, Y.-P. Hung, and H. Chu, "Mug-tree: a playful mug to encourage healthy habit of drinking fluid regularly," *the Late Breaking Results (LBR) session of Proc. UBICOMP*, 2007.

[33] Fitbit. (2015) Fitbit official site for activity trackers and more. [Online]. Available: https://www.fitbit.com

[34] U. B. Jawbone. (2015) A smarter fitness tracker for a fitter you. [Online]. Available: https://jawbone.com/up

[35] Misfit. (2015) Premium fitness + sleep monitor. [Online]. Available: http://misfit.com/products/shine

[36] S. Purpura, V. Schwanda, K. Williams, W. Stubler, and P. Sengers, "Fit4life: the design of a persuasive technology promoting healthy behavior and ideal weight," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 423–432.

[37] Spire. (2015) The mindfulness and activity tracker. [Online]. Available: https://www.spire.io/?

[38] T. Fritz, E. M. Huang, G. C. Murphy, and T. Zimmermann, "Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 487–496.

[39] D. Ledger and D. McCaffrey, "Inside wearables: How the science of human behavior change offers the secret to long-term engagement," *Endeavour Partners*, 2014.

[40] J. Kopp, "Self-monitoring: A literature review of research and practice," in *Social Work Research and Abstracts*, vol. 24, no. 4. Oxford University Press, 1988, pp. 8–20.

[41] L. E. Burke, J. Wang, and M. A. Sevick, "Self-monitoring in weight loss: a systematic review of the literature," *Journal of the American Dietetic Association*, vol. 111, no. 1, pp. 92–102, 2011.

[42] S. Fox and M. Duggan, *Tracking for health*. Pew Research Center's Internet & American Life Project, 2013.

[43] W. Randolph and K. Viswanath, "Lessons learned from public health mass media campaigns: Marketing health in a crowded media world*," *Annu. Rev. Public Health*, vol. 25, pp. 419–437, 2004.

[44] P. A. Keller and D. R. Lehmann, "Designing effective health communications: a meta-analysis," *Journal of Public Policy & Marketing*, vol. 27, no. 2, pp. 117–130, 2008.

[45] L. T. Cowan, S. A. Van Wagenen, B. A. Brown, R. J. Hedin, Y. Seino-Stephan, P. C. Hall, and J. H. West, "Apps of steel: are exercise apps providing consumers with realistic expectations? a content analysis of exercise apps for presence of behavior change theory," *Health Education & Behavior*, p. 1090198112452126, 2012.

[46] A. Kadomura, C.-Y. Li, Y.-C. Chen, H.-H. Chu, K. Tsukada, and I. Siio, "Sensing fork and persuasive game for improving eating behavior," in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 2013, pp. 71–74.

[47] J.-L. Lo, T.-Y. Lin, H.-H. Chu, H.-C. Chou, J.-H. Chen, J. Y.-J. Hsu, and P. Huang, *Playful tray: adopting Ubicomp and persuasive techniques into play-based occupational therapy for reducing poor eating behavior in young children.* Springer, 2007.

[48] O. Amft and G. Tr, "On-body sensing solutions for automatic dietary monitoring," *IEEE pervasive computing*, no. 2, pp. 62–70, 2009.

[49] M. Garhammer, "Pace of life and enjoyment of life," *Journal of Happiness Studies*, vol. 3, no. 3, pp. 217–256, 2002.

[50] Gizmodo. (2015) 7 apps and sites to track your time in 2015. [Online]. Available: http://fieldguide.gizmodo.com/7-apps-and-sites-to-track-your-time-in-2015-1677057680

[51] F. Zini, M. Reinstadler, and F. Ricci, "Life-logs aggregation for quality of life monitoring," in *Proceedings of the 5th International Conference on Digital Health 2015.* ACM, 2015, pp. 131–132.

[52] M. Al Masum Shaikh, K. I. Molla, and K. Hirose, "Automatic life-logging: A novel approach to sense real-world activities by environmental sound cues and common sense," in *Computer and Information Technology, 2008. ICCIT 2008. 11th International Conference on.* IEEE, 2008, pp. 294–299.

[53] J. Braverman, "Testimonials versus informational persuasive messages: The moderating effect of delivery mode and personal involvement," *Communication Research*, 2008.

[54] J.-F. Coget, Y. Yamauchi, and M. Suman, "The internet, social networks and loneliness," *It & Society*, vol. 1, no. 1, p. 180, 2002.

[55] J. Holt-Lunstad, T. B. Smith, and J. B. Layton, "Social relationships and mortality risk: a meta-analytic review," *PLoS Med*, vol. 7, no. 7, p. e1000316, 2010.

[56] J. T. Cacioppo and W. Patrick, *Loneliness: Human nature and the need for social connection.* WW Norton & Company, 2008.

[57] Narrative. (2015) Narrative clip - the wearable camera for moments that matter. [Online]. Available: http://getnarrative.com

[58] W. James, *The principles of psychology.* Read Books Ltd, 2013.

[59] E. Hatfield, J. T. Cacioppo, and R. L. Rapson, *Emotional contagion.* Cambridge university press, 1994.

[60] J. Hernandez, M. E. Hoque, W. Drevo, and R. W. Picard, "Mood meter: counting smiles in the wild," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing.* ACM, 2012, pp. 301–310.

[61] H. Tsujita and J. Rekimoto, "Happinesscounter: smile-encouraging appliance to increase positive mood," in *CHI'11 Extended Abstracts on Human Factors in Computing Systems.* ACM, 2011, pp. 117–126.

[62] Face++. (2015) Face++ - leading face recognition on cloud. [Online]. Available: http://www.faceplusplus.com

[63] R. KORMI-NOURI, L.-G. NILSSON, and N. Ohta, "The novelty effect: Support for the novelty-encoding hypothesis," *Scandinavian journal of psychology*, vol. 46, no. 2, pp. 133–143, 2005.

[64] D. L. Thomas and E. Diener, "Memory accuracy in the recall of emotions." *Journal of Personality and Social Psychology*, vol. 59, no. 2, p. 291, 1990.

[65] D. A. Small and N. M. Verrochi, "The face of need: Facial emotion expression on charity advertisements," *Journal of Marketing Research*, vol. 46, no. 6, pp. 777–787, 2009.

[66] P. C. Watkins, D. L. Grimm, and R. Kolts, "Counting your blessings: Positive memories among grateful persons," *Current Psychology*, vol. 23, no. 1, pp. 52–67, 2004.

[67] L. B. Andersen, P. Schnohr, M. Schroll, and H. O. Hein, "All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work," *Archives of internal medicine*, vol. 160, no. 11, pp. 1621–1628, 2000.

[68] J. A. Berlin and G. A. Colditz, "A meta-analysis of physical activity in the prevention of coronary heart disease," *American journal of epidemiology*, vol. 132, no. 4, pp. 612–628, 1990.

[69] E. G. Wilmot, C. L. Edwardson, F. A. Achana, M. J. Davies, T. Gorely, L. J. Gray, K. Khunti, T. Yates, and S. J. Biddle, "Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis," *Diabetologia*, vol. 55, pp. 2895–2905, 2012.

[70] R. S. Paffenbarger Jr and W. E. Hale, "Work activity and coronary heart mortality," *New England Journal of Medicine*, vol. 292, no. 11, pp. 545–550, 1975.

[71] D. John, "Science of workplace fitness," November 2014, mIT Media Lab Tools for Wellbeing lecture series, conducted from Cambridge, MA.

[72] L. C. Hawkley and J. T. Cacioppo, "Loneliness matters: a theoretical and empirical review of consequences and mechanisms," *Annals of Behavioral Medicine*, vol. 40, no. 2, pp. 218–227, 2010.

[73] V. S. Conn, A. R. Hafdahl, P. S. Cooper, L. M. Brown, and S. L. Lusk, "Meta-analysis of workplace physical activity interventions," *American journal of preventive medicine*, vol. 37, no. 4, pp. 330–339, 2009.

[74] R. K. Dishman, "Compliance/adherence in health-related exercise." *Health Psychology*, vol. 1, no. 3, p. 237, 1982.

[75] K. C. Harris, L. K. Kuramoto, M. Schulzer, and J. E. Retallack, "Effect of school-based physical activity interventions on body mass index in children: a meta-analysis," *Canadian Medical Association Journal*, vol. 180, no. 7, pp. 719–726, 2009.

[76] Y. Netz, M.-J. Wu, B. J. Becker, and G. Tenenbaum, "Physical activity and psychological well-being in advanced age: a meta-analysis of intervention studies." *Psychology and aging*, vol. 20, no. 2, p. 272, 2005.

[77] A. V. Carron, W. N. Widmeyer, and L. R. Brawley, "Group cohesion and individual adherence to physical activity." *Journal of Sport & Exercise Psychology*, 1988.

[78] D. M. Bravata, C. Smith-Spangler, V. Sundaram, A. L. Gienger, N. Lin, R. Lewis, C. D. Stave, I. Olkin, and J. R. Sirard, "Using pedometers to increase physical activity and improve health: a systematic review," *Jama*, vol. 298, no. 19, pp. 2296–2304, 2007.

[79] C. B. Chan, D. A. Ryan, and C. Tudor-Locke, "Health benefits of a pedometer-based physical activity intervention in sedentary workers," *Preventive medicine*, vol. 39, no. 6, pp. 1215–1222, 2004.

[80] R. A. Findlay, "Interventions to reduce social isolation amongst older people: where is the evidence?" *Ageing and Society*, vol. 23, no. 05, pp. 647–658, 2003.

[81] M. Cattan, M. White, J. Bond, and A. Learmouth, "Preventing social isolation and loneliness among older people: a systematic review of health promotion interventions," *Ageing and society*, vol. 25, no. 01, pp. 41–67, 2005.

[82] A. Stibe, H. Oinas-Kukkonen, and T. Lehto, "Exploring social influence on customer engagement: a pilot study on the effects of social learning, social comparison, and normative influence," in *System Sciences (HICSS)*. IEEE, 2013, pp. 2735–2744.

[83] B. J. Fogg, "A behavior model for persuasive design," in *Persuasive Technology*. ACM, 2009, p. 40.

[84] A. Pentland, *Social Physics: How Good Ideas Spread-The Lessons from a New Science*. Penguin, 2014.

[85] R. Cialdini, "Science of persuasion," November 2012. [Online]. Available: https://www.youtube.com/watch?v=cFdCzN7RYbw

[86] A. R. Hollenbeck, "Problems of reliability in observational research," *Observing behavior*, vol. 2, pp. 79–98, 1978.

[87] C. Tudor-Locke, S. A. Ham, C. A. Macera, B. E. Ainsworth, K. A. Kirtland, J. P. Reis, and C. D. Kimsey Jr, "Descriptive epidemiology of pedometer-determined physical activity." *Medicine and Science in Sports and exercise*, vol. 36, no. 9, pp. 1567–1573, 2004.

[88] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.

[89] N. A. Christakis and J. H. Fowler, "The spread of obesity in a large social network over 32 years," *New England journal of medicine*, vol. 357, no. 4, pp. 370–379, 2007.

[90] B. J. Rhodes and P. Maes, "Just-in-time information retrieval agents," *IBM Systems journal*, vol. 39, no. 3.4, pp. 685–704, 2000.

[91] R. Hansson, P. Ljungstrand, and J. Redström, "Subtle and public notification cues for mobile devices," in *Ubicomp 2001: Ubiquitous Computing.* Springer, 2001, pp. 240–246.

[92] R. Vertegaal, C. Dickie, C. Sohn, and M. Flickner, "Designing attentive cell phone using wearable eyecontact sensors," in *CHI'02 extended abstracts on Human factors in computing systems.* ACM, 2002, pp. 646–647.

[93] P. Tarasewich, T. Bhimdi, and M. Dideles, "Increasing the effectiveness of notification cues in mobile environments," *AMCIS 2004 Proceedings*, p. 402, 2004.

[94] G. Glass. (2016) Google glass help. [Online]. Available: https://support.google.com/glass/answer/3064128?hl=en

[95] Wearscript. (2015) Wearscript: Js with batteries included for glass ? wearscript .0.1.0 documentation. [Online]. Available: http://www.wearscript.com/en/latest/

[96] A. Oulasvirta, S. Tamminen, V. Roto, and J. Kuorelahti, "Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci," in *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 2005, pp. 919–928.

[97] N. Győrbíró, Á. Fábián, and G. Hományi, "An activity recognition system for mobile phones," *Mobile Networks and Applications*, vol. 14, no. 1, pp. 82–91, 2009.

[98] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[99] B. Gardner, L. Smith, F. Lorencatto, M. Hamer, and S. J. Biddle, "How to reduce sitting time? a review of behaviour change strategies used in sedentary behaviour reduction interventions among adults," *Health psychology review*, pp. 1–24, 2015.

[100] J. Maitland, K. Siek, and M. Chalmers, "Persuasion not required: Obstacles faced by low-income caregivers to improve dietary behaviour," in *3rd International Conference on Pervasive Computing Technologies for Healthcare*, vol. 2009, 2009, pp. 1–8.

[101] J. McGonigal, *Reality is broken: Why games make us better and how they can change the world.* Penguin, 2011.

[102] M. Pinquart and S. Sörensen, "Influences of socioeconomic status, social network, and competence on subjective well-being in later life: a meta-analysis." *Psychology and aging*, vol. 15, no. 2, p. 187, 2000.