Aalto University

School of Science

Master’s Programme in International Design Business Management

Designing for Persuasion

An Experimental Case Study: Oura Ring

Verneri Jäämuru

Master’s Thesis

September 30th, 2019

Supervisor: **Matti Vartiainen,** Professor

Advisor: **Teemu Kurppa**, M.Sc. Computer Science

Aalto University
School of Science

Abstract of master's thesis

Author Verneri Jäämuru

Title of thesis Designing for Persuasion, An Experimental Case Study: Oura Ring

Degree Master of Science

Degree programme International Design Business Management

Thesis advisor Teemu Kurppa

Year of approval 2019 **Number of pages** 70 **Language** English

Abstract

The information age has ushered in an era of digital devices galore, and swaths of data to boot. These devices and the ease of collecting data on user behavior have contributed to the growing interest in persuasive design, the application of psychological and social principles of influence to elicit a desired action from the user of a product or service.

This Master's Thesis explores the potential of persuasive design to induce healthy behavioral change in the context of the sleep and activity tracker Oura ring and its adjoined mobile application. In partnership with Oura Health, a persuasive design is developed and tested against a control. The persuasive design incorporates theories of narrative persuasion and seeks to enhance the onboarding process of users into the Oura app. In our experiment, we focus on message timing and format, and postulate that a more narrative like onboarding process will be easier to digest and absorb, leading to increased engagement and a possible uptick in the target behavior of ring wear.

Compared to control, the engagement metrics of time in app and sessions per day were observed to increase in the treatment group, although not statistically significantly. Additionally, the rate of ring wear was observed to increase, but without significance. Interestingly, among the subset of users found to have been traveling internationally, the rate of ring wear increased by 3.4% ($P < 0.037$).

Despite the lack of significance, our research finds that persuasive design, and narrative persuasion specifically, hold promise for the design of digital products aimed at promoting healthy habits. In order to achieve the desired outcomes, iteration of persuasive designs is advised. Additionally, engagement metrics such as time in app or sessions per day, show promise as leading indicators for behavioral change.

Keywords persuasive design, behavioral change, habits, user experience, health, wellness, narrative persuasion, product design

Contents

List of Figures

List of Tables

List of Abbreviations

CLT	Cognitive Load Theory
ELM	Elaboration Likelihood Model
FBM	Fogg Behavior Model
FTUE	First Time User Experience
IT	Information Technology
KPI	Key Performance Indicator
QS	Quantified Self
RCT	Randomized Controlled Trial
ROI	Return on Investment
UI	User Interface
UX	User Experience

Chapter 1

Introduction

> *Every reader, as he reads, is actually the reader of himself. The writer's work is only a kind of optical instrument he provides the reader so he can discern what he might never have seen in himself without this book.*
>
> \- Marcel Proust, 1933

Although Proust is referring to written literature, the essence of persuasion in general is captured well by Proust's remark above. Persuasion should not be coercive nor shall it strip a person of their autonomy. When exercised properly persuasion shall only be a means by which to help another person reach an enlightened state of mind. There is huge potential behind persuasive design and technology to act as a benevolent force in the world but the importance of ethical decision making can not be understated. With that in mind, this Master's Thesis explores the potential of persuasive design to induce healthy behavioral changes in the context of the sleep and activity tracker Oura ring and its adjoined mobile application.

This chapter gives the background on the topic of persuasive design, explicates the context and nature of the research problem, and lastly reviews the outline of this research paper as a whole.

1.1 Background

Throughout the ages humans have sought to understand the inner workings of the mind. As early as the 5th century B.C. ancient civilizations around the world engaged in philosophical studies of psychology. In the modern world, our curiosity for understanding the mind is unchanged but our tools for doing so have gotten much better. The information age brought along with it computers, the internet, mobile devices, and swaths of data to boot. These devices and the ease of collecting data on user behavior have contributed to the growing interest in persuasive design, the application of psychological and social principles of influence to elicit a desired action from the user of a product or service.

Persuasive design allows the creator of a product or a piece of technology to influence the behavior of its user in a way that benefits both the user and the creator. Persuasive design is often used in marketing, organizational management, sales, and in many health-related applications. The methods by which persuasion and influence can be used to impact human behavior are of high interest and significance to many industries because they have the potential to grow sales, increase efficiency, and in the best cases prevent illness or even injury.

Narrative persuasion is just one of many persuasive design methods, but perhaps one of the most ubiquitous methods thanks to the relative ease of its implementation. Narratives are effective methods of persuasion in that they absorb and engage the reader into a plot, ultimately leading to a process of reflection. In this reflection, the reader compares and adjusts personal mental models to become more aligned and consistent with beliefs from the narrative (Hamby, Brinberg, & Jaccard, 2018). Persuading in the form of narratives or stories is hardly a novel concept, but its application in mobile user experiences is academically largely unexplored.

The digital nature of a mobile UX renders it to be an interesting medium for the conveyance of narratives. Narratives contained inside of mobile applications can be presented in a variety of different ways, with a mixture of verbal, visual, or even audible elements. Furthermore these narrative elements, thanks to their digital nature, can relatively easily be created, tested, and improved. These aspects position mobile UX to be a prime platform for the application of not only narratives but of persuasive design in general.

1.2 Research Objectives and Questions

This research seeks to understand how persuasive design can be applied to the Oura ring mobile application to induce healthy behavioral change amongst users of the application. Oura Health is a company that sells a sleep and activity tracker called Oura ring. At the core of Oura Health's value proposition is a mission to empower people to own their full potential. Oura does this primarily through Oura ring, a wearable computer in the form of a ring, and the Oura mobile application. The Oura ring tracks sleep and activity levels which then, along with insights, are presented in the Oura mobile application. Insights are focused on presenting novel information regarding the user's sleep or activity, and giving actionable steps that the user can take to improve wellbeing.

Oura can show and explain to users the path to better sleep and health, such as implementing a consistent bedtime, but ultimately it is up to the user to implement certain lifestyle changes to improve wellbeing. Thus, by enabling users to more effectively adopt healthier lifestyle habits Oura can create more value for its customers. And therein lies the motivation for this research case, to better understand how persuasive design could be used to help Oura users more easily adopt healthy behaviors. Furthermore, the primary objective of this thesis is to productize findings from theoretical research into the Oura mobile application and quantitatively measure any changes in user behavior.

Given the motivation, objective, and context of this research, there are three questions specifically that we seek to answer:

> *1) How can persuasive design be used to induce healthy behavioral change?*
>
> *2) Can narrative persuasion specifically, be used to increase engagement rate in the context of a mobile application?*
>
> *3) Can message timing and format be altered to increase engagement and the rate of ring wear among iOS users of the Oura application?*

There is a swathe of academic literature surrounding the research questions stated above. Influence, persuasive design, and behavioral change have all been widely studied in the last few decades by authors such as Cialdini, B.J. Fogg, and Oinas-Kukkonen. More recently narratives and storytelling have been explored by authors such as Hamby, Brinberg, and Jaccard. While these concepts have been studied extensively in regards to increasing health related behaviors, this research contributes by exploring narrative persuasion specifically in the context of mobile applications.

This experimental case study seeks to answer the above mentioned research questions by first conducting a thorough review of prior literature. Based on the existing research, a hypothesis is formed and a conceptual framework is constructed. This hypothesis and adjoining conceptual framework is then tested in vivo to determine if and how persuasive design can be used to induce healthy behavioral change in the context of the Oura mobile application.

1.3 Structure of the Thesis

In the following chapter, the theoretical background of this research is explored. Relevant literature regarding wearable technology, behavioral change, persuasive design, and narratives in UX design will be reviewed. Lastly, theoretical findings are synthesized and a specified hypothesis is postulated.

In Chapter 3, the methodology of this experimental case study is described. The research questions are reviewed, the research approach is stated, and the process of productizing the theoretical findings is explained in detail. Lastly, the process of data collection and analysis is discussed.

The final two chapters of this thesis are spent describing and discussing the findings of the quantitative experiment. The collected data is reviewed, interpreted, and both practical and theoretical contributions are expressed. And lastly the study is evaluated and further directions of research are suggested.

Chapter 2

Theoretical Background

Our research is grounded in theoretical background that spans the disciplines of psychology, technology, and data science. In this chapter, we review previously researched topics and concepts to which this research paper is attached. Firstly, we review the underlying psychological phenomenon behind persuasion. Secondly, we dive into the applications of persuasion in a technological context. After that, we examine the efficacy of narratives to persuade, followed by a closer look at how user behavior can be measured in a UX context. Lastly, we re-examine our research questions and synthesize findings from theoretical research in order to address them, arriving at a clarified and testable hypothesis.

2.1 Psychology of Persuasion

To persuade is to be human. Throughout time, humans have engaged in persuasion through rhetoric, or spoken word. Torning and Oinas-Kukkonen (2009) said it well, “whenever we communicate deliberately with a clear purpose and outcome in mind, we are engaging in persuasion.” Already in 350 B.C., Aristotle defined rhetoric as “…the faculty of observing in any given case the available means of persuasion” (Prt 2).

It is clear that persuasion has ancient roots, but only in the last century, however, has humanity begun to scientifically explain the phenomenon. Much of the research has revolved around psychological principles with a primary focus on human cognition, motivations, and behaviors. The following sections explore this prior research.

2.1.1 Elaboration Likelihood Model

According to Petty and Cacioppo (1986), in the first half of the 20th century research relating to persuasion focused mostly on studying human attitudes.

Researchers however, had difficulty using attitudes to predict behavior. By the 1970's, researchers such as Ajzen & Fishbein (1977) and Fazio, Zanna, & Cooper (1978) identified instances where attitudes would and would not predict behavior. Gradually, towards the end of the century, research progressed to focus on understanding attitude change and furthermore understanding the interplay between attitudes and behaviors.

Following these discoveries, Petty and Cacioppo established the Elaboration Likelihood Model (1986). The ELM posits two different cognitive routes to

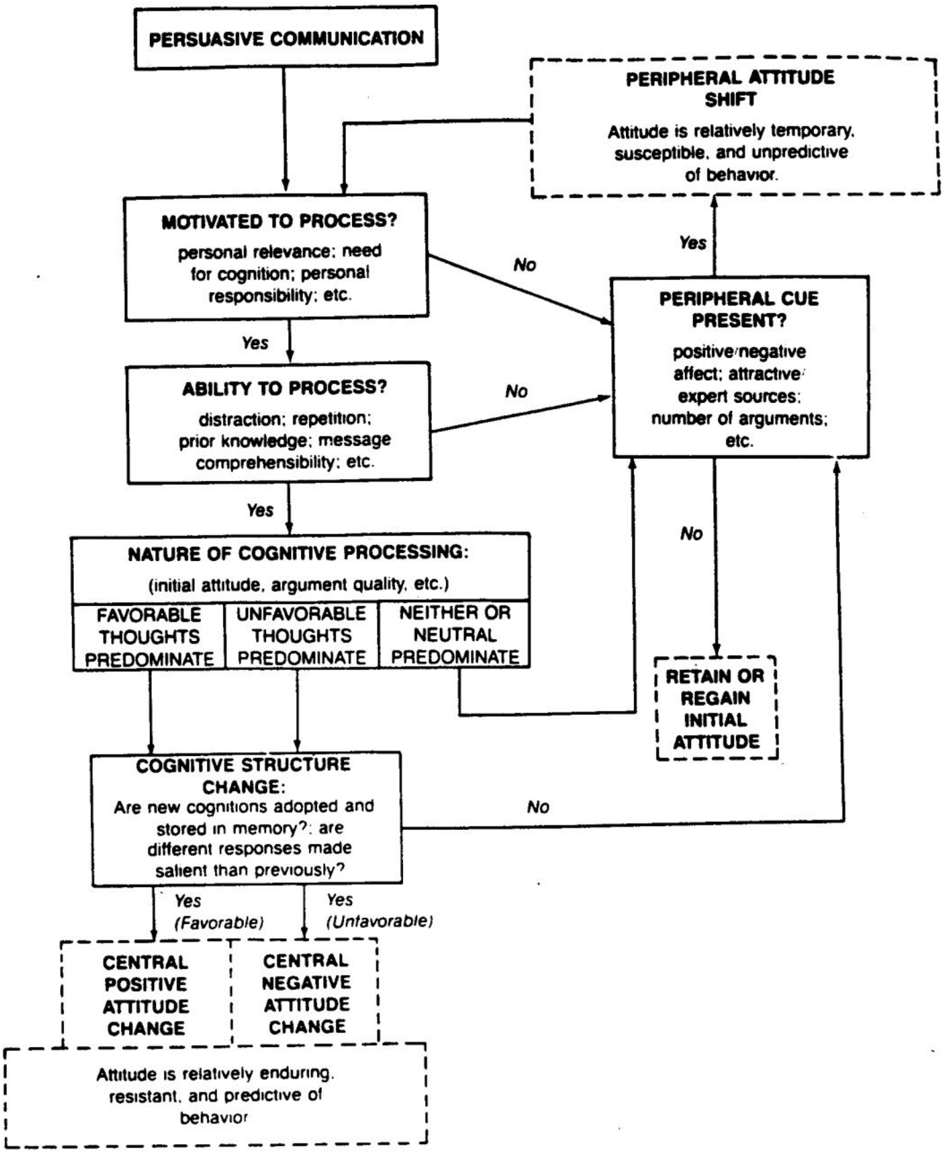


Figure 1. Central and peripheral routes to persuasion (Petty & Cacioppo, 1986, p. 126).

Persuasion (Figure 1). The central route, which results from the subject's careful consideration of the true merits of the presented persuasive argument. And the peripheral route, which results from simple cues in the context of the environment in which the persuasion is occuring. Note that in the peripheral route the scrutiny of processing the true merits of the presented information does not occur; for example, a person may come to like a political candidate because the music in a recently heard commercial created a general feeling of pleasantness (Petty & Cacioppo, 1986).

The ELM as shown in Figure 1 creates a general framework for organizing and understanding the underlying processes behind the effectiveness of persuasive communication. For the purposes of our research, it is important to note that attitude changes stemming from the central route results in new attitudes that are relatively enduring, resistant, and predictive of behavior.

2.1.2 Fogg Behavior Model

While the ELM primarily focuses on understanding attitude change, the Fogg Behavior Model (FBM) takes things further and attempts to explain when and why an actual instance of behavior occurs. Created by Stanford Persuasive Technology Lab's BJ Fogg, the FBM asserts that for a person to perform a target behavior, she must be (1) sufficiently motivated, (2) have the ability to perform the behavior, and (3) be triggered to perform the behavior (Fogg, 2009a). For a

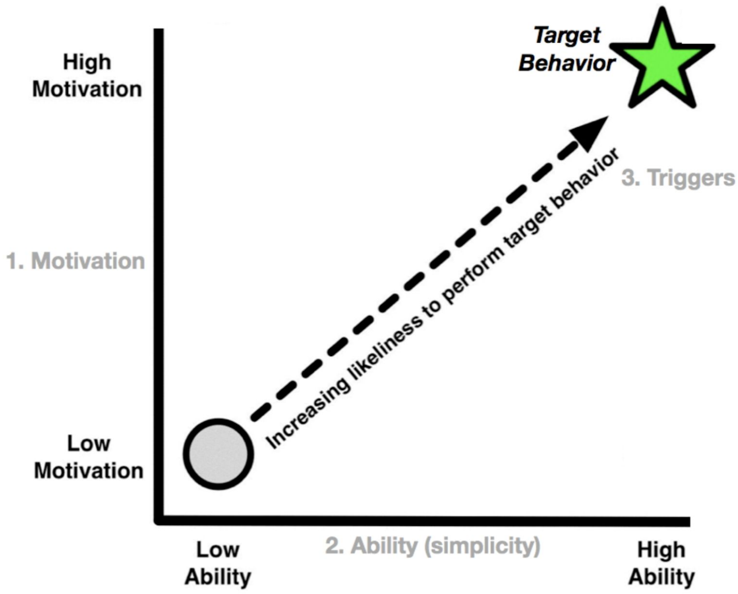


Figure 2. The Fogg Behavior Model and its three factors: motivation, ability, and triggers (Fogg, 2009a, p.2).

desired target behavior to take place, all three factors must take place at the same time. Figure 2 displays the tradeoff between motivation and ability. If both motivation and ability to perform a behavior are high, given a trigger, the likelihood of that behavior occurring is also high. Note that for there to be any chance of a given behavior occurring both motivation and ability must be non-zero. This model gives a clear way of thinking about behavior change and how to begin to design it.

Fogg elaborates on the FBM by identifying common sources of motivation and prevailing elements of simplicity (ability). These common motivations and elements of simplicity can be used by designers to more effectively design for persuasion. In the FBM, there are three core motivators, each of which has two sides: *pleasure/pain*, *hope/fear*, and *social acceptance/rejection*. The six elements of simplicity are: *time*, *money*, *physical effort*, *brain cycles*, *social deviance*, and *non-routine*. For example, I might be motivated to mow my lawn due to the pleasure of having a tidy yard or the fear of social scorn if it is left unkempt. Furthermore, if I do not have time to mow the lawn my ability to perform the behavior is low and the job will not get done. If, however, I am highly motivated, I might seek someone to hire for the job. In that case, if I have the money, my ability to perform the target behavior is once again within reach and it is likely that I may choose to do so.

In essence, understanding the driving motivation can explain a person's enthusiasm to perform a target behavior. Each element of simplicity on the other hand begins to paint a picture of how complex or simple it will be for a person to complete the behavior. In the context of designing for persuasion, an effort is typically focused on channeling a certain motivation and increasing ability via reducing complexity. Increasing innate ability is often problematic as humans are by nature lazy and very reluctant to exert effort in learning new skills.

2.1.3 Cognitive Load Theory

An especially important theory surrounding the design of persuasion is the Cognitive Load Theory (CLT) established by John Sweller. The CLT was born from Sweller's study on cognitive load during problem solving and its effects on learning in 1988. Cognitive load simply refers to the amount of working memory resources used during problem solving or task completion. Sweller (1988) identified that the difference between expert and novice problem solvers was that experts were able to work from memorized schema, or in other words, memorized patterns of thought or behavior that organize information. Furthermore, Sweller stated that the way in which a problem was presented affects the potential learning or formation of schema by way of placing a heavier load on working memory.

These findings paved the way for the formation of CLT, which differentiates cognitive load into three types: intrinsic, extraneous, and germane (Sweller, Merrienboer, & Paas, 1998). Intrinsic load is load inherent to the given task, extraneous load is load stemming from the way or manner that the task is presented, and germane load is load caused by the formation of new schemas. To maximize performance on a task, the design of the task should then strive to decrease extraneous load. If the task is to be performed again in the future, it would be well served for the design of the task to also increase germane cognitive load.

In the context of our experimental case study, CLT gives us insight into how different content might be processed by a user. According to CLT, in designing a new UX, it is best to utilize any existing schema a user may have, and if novel ideas are to be presented, they should be done so in a manner that removes any extra load on cognition. It has indeed been shown that the more cognitive load is reduced by cues, the better the retention and absorption of multimedia learning (Xie et al., 2017).

2.1.4 Methods of Persuasion

Having reviewed some of the most central psychological frameworks and models for understanding behavior, we are now free to explore some of the more effective methods of persuasion. It is good to keep in mind that at best we can only design for persuasion. Ultimately, persuasion occurs internally and an observer must persuade herself (either consciously or subconsciously) of the validity of the message being conveyed (Geddes, 2016). Many attempts at persuasion fail. We can increase our odds of success, however, by further studying general psychological concepts of persuasion outlined in the following paragraphs.

> *There is no expedient to which a man will not resort to avoid the real labor of thinking.*
>
> \- Robert Cialdini, 1984

Much of the modern psychological literature on persuasion tracks back to Robert B. Cialdini. Cialdini was no stranger to the concept of cognitive load, as the quote above illustrates, but his primary area of research was influence and persuasion. In 1984, he outlined six of what he believed to be the most effective methods of persuasion: *Reciprocation*, *Commitment and Consistency, Social Proof, Liking, Authority, and Scarcity.* Many of these persuasion methods or tactics are particularly applicable in sales and marketing contexts, however, the underlying psychology behind them applies to many everyday situations. Although not particularly rooted in technological contexts, for many still to this day, Cialdini's methods form the backbone for approaching the design of persuasive products, services, or marketing.

Many other methods of persuasion exist and often the lines between different methods are quite blurred and overlapping. In academic literature, goal-setting for instance has been widely studied. In 1991, Locke and Latham retorted that conscious human behavior is by nature purposeful, and to increase, for example, an employee's performance, a manager could set for that employee a moderately challenging task or goal. This goal setting theory is, for instance, very similar to Cialdini's commitment and consistency principle, simply approached from and studied in a slightly different context.

Furthermore, the rise of digital technologies such as smartphones, and personal computers, has given birth to a new era in which many other niches and sub-niches of psychological persuasion have flourished. Cialdini's persuasion method of liking, for instance, has to a large extent taken the form of personalization in the modern age. The more personalized an experience, the more likely a person is to like that experience, resulting in that person becoming more susceptible to persuasion. Additionally, the better a system fits a person's needs and requirements, the more likely that person is to build trust with the system, potentially leading to more persuasion in the future (Berkovsky, Freyne, & Oinas-Kukkonen, 2012).

Countless other methods and tactics for persuasion exist still. As stated above, however, it is good to be cognizant of the realization that many of these methods overlap considerably, and that the core psychological principles outlined above serve us well in describing and understanding the majority of them. Of the other methods explored previously in academic research, many build upon the aforementioned core principles and drill deeper into a given context or use case. One such method, narrative persuasion, will be explored in greater detail in section 2.3.

2.2 Persuasive Technology

For millennia, persuasion was limited to spoken word. To recap, persuasion is commonly referred to as communicating deliberately with a clear purpose and outcome in mind, and although persuasion is not a new phenomenon, the building of machines and technology to do the persuading on our behalf is (Torning & Oinas-Kukkonen, 2009). In this section, we review prior literature on persuasive technology, or in other words, the use of technology as a vehicle for persuasion. First, we look into the methodology behind persuasive systems design, then review procedures on creating persuasive technology. Following, we touch on applications of persuasive technology pertaining to our research, and lastly we consider the ethical aspects of persuasive technology creation.

2.2.1 Persuasive System Design

While persuasive technology is defined as interactive information technology designed for changing users' attitudes or behavior (Fogg, 2003), persuasive systems dive a bit deeper and are more specifically defined as "computerized software or information systems designed to reinforce, change or shape attitudes or behaviors" (Oinas-Kukkonen & Harjumaa, 2008, p. 202). The difference is subtle, but persuasive systems can be thought of as a type of persuasive technology. Established by Oinas-Kukkonen and Harjumaa, the persuasive systems model provides a framework from which to think of and approach persuasive technology creation.

Oinas-Kukkonen and Harjumaa established seven postulates for persuasive systems, that according to them, need to be addressed when designing persuasive systems. The postulates are shown in Table 2.

Table 1. Postulates behind persuasive systems (Oinas-Kukkonen & Harjumaa, 2009, p. 487).

1. Information technology is never neutral.
2. People like their views about the world to be organized and consistent.
3. Direct and indirect routes are key persuasion strategies.
4. Persuasion is often incremental.
5. Persuasion through persuasive systems should always be open.
6. Persuasive systems should aim at unobtrusiveness.
7. Persuasive systems should aim at being both useful and easy to use.

These postulates address points regarding the end user, persuasion strategies, and actual persuasive system features that help us move forward in designing actual persuasive systems.

Another crucial aspect to consider when designing persuasive systems is the persuasion context. This typically requires a thorough understanding of what happens during persuasion, namely understanding the roles of a persuader, a persuadee, a message, a channel, and the broader persuasion context (Oinas-Kukkonen & Harjumaa, 2008). To assist in analyzing persuasion context Oinas-Kukkonen & Harjumaa created Figure 3.

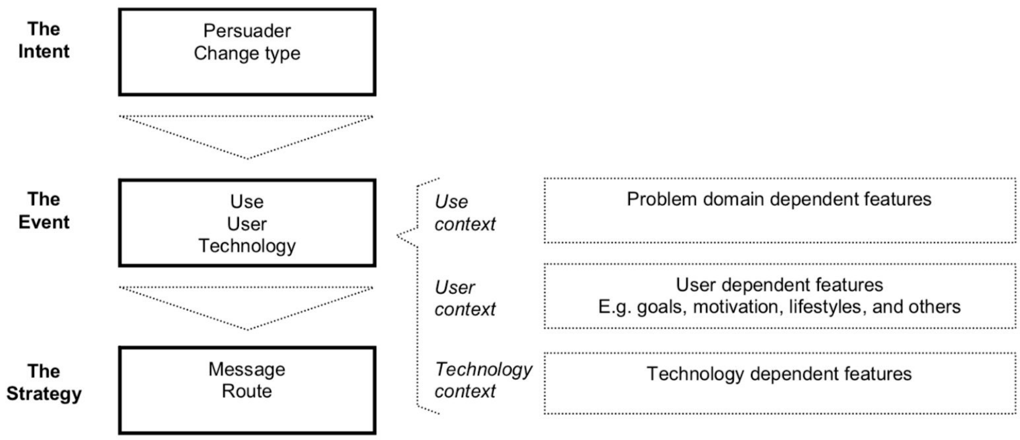


Figure 3. Analyzing the persuasion context (Oinas-Kukkonen & Harjumaa, 2009, p. 489).

As Figure 3 shows, attention should be paid to the intent of the persuader, the persuasion event itself, and the persuasion strategy. Furthermore, the context of the persuasion event itself can be broken down into the use context, the user context, and the technology context. This figure and the postulates mentioned above form a solid basis from which to theoretically approach the creation of persuasive systems and technology.

2.2.2 Creating Persuasive Technology

Drawing from demonstrated success in industry practice, Fogg has constructed an eight step design process for the creation of persuasive technology. Most projects aimed at creating persuasive technology fail due to unrealistic expectations. Fogg encourages designers to fail fast with attempts at persuasion and stresses the importance of learning how to succeed with persuasive design projects in the early stages; getting small victories at first and then moving on from there (Fogg, 2009b). The steps are outlined below:

Step 1: Choose a simple behavior to target

Here the design team should select the smallest, simplest, behavior that still matters. Large vague goals should be broken down by selecting a smaller simpler goal that adequately approximates the larger goal or by selecting a smaller goal that acts as a first step towards the larger goal. The first step is the most important step in designing a new persuasive technology.

Step 2: Choose a receptive audience

If the design team has a choice in selecting the audience, the most receptive audience should be chosen, e.g., if the goal is to get people to eat healthier, the team should select an audience that has already demonstrated a desire to eat healthier. Secondly, the team should consider how familiar the target audience is with using the technology channel. Early adopters are typically a good initial audience. Note that in some cases Step 2 may be completed before Step 1, or Step 2 may force the team to revisit Step 1.

Step 3: Find what prevents the target behavior

What is preventing the desired behavior from being exhibited? The answer typically falls into a combination of the following three categories: lack of motivation, lack of ability, or the lack of a well-timed trigger to perform the behavior. A thorough investigation is best, as the core of the persuasive technology should be designed to either boost motivation, increase ability, or introduce a trigger accordingly. Note that if motivation and ability are lacking it is advised to reconsider the previous steps.

Step 4: Choose a familiar technology channel

All of the first three steps should be considered when deciding on a technology channel. The target behavior, the audience, and whatever is preventing the behavior from occurring will all largely impact whether a given channel is successful or not. Most importantly, however, the channel should be familiar to the target audience. If the channel is new to the audience, it is unrealistic to expect the audience to learn the new channel and to adopt a new behavior change on top of that.

Step 5: Find relevant examples of persuasive technology

In this step the design team should seek out relevant prior persuasive technology that has been deemed to be successful. Fogg advocates for the examination of at least nine case examples: three that achieve a similar behavior, three that target a similar audience, and three that use the same technology channel that the design team is intending to use.

Step 6: Imitate successful examples

From the benchmarks identified in Step 5, the effective strategies should be imitated. There is no shame in imitating pre-existing solutions, as identifying what works in other cases and successfully replicating it can be surprisingly difficult.

Step 7: Test and iterate quickly

The design team should then test and iterate persuasive experiences as quickly as possible. Ideally, testing cycles should be no longer than 10 hours, allowing the team to quickly see what works and what does not. These tests should not be scientific and focus should be paid on how people react, and in particular measuring their behavior.

Step 8: Expand on success

If a behavior change has been observed, no matter how small, the team can then begin to expand on its success. Expanding on success should be systematic. Only one, or at most two, attributes from the successful case should be improved upon at a time.

The strength of Fogg's process lies in its practicality. While the persuasive system design methodology reviewed in section 2.2.1 offers a grounded theoretical approach, Fogg's eight step process provides more concrete and actionable steps to take in order to create persuasive technology.

2.2.3 Applications of Persuasive Technology

Next, we review academic literature on applications of persuasive technology. Applications of persuasive technology are of course vast, but for the purposes of this research we focus on applications centered around health promotion and activity tracking. A few choice examples are outlined below.

Halko and Kientz (2010) studied the relationship between personality type and preferred persuasion strategy in health-promoting mobile applications, finding several correlations between the two. In terms of strategies the researchers specifically looked at instruction style, social feedback, motivation type, and reinforcement type. Unfortunately, their research was not conducted in vivo; study participants were simply shown storyboards of how a mobile application might look and then prompted to fill out surveys indicating their preferences, followed by a Big Five Inventory to assess their personality type. The study, however, succeeds in illustrating that a one-size-fits-all approach to persuasive technology could potentially be improved by tailoring the persuasive technology to fit the needs of the user, ultimately resulting in higher success rates.

Consolvo et al. (2009), however, studied a type of mobile persuasive technology in vivo, to explore the efficacy of different goal setting methods to encourage physical activity. In their study, 28 individuals used an UbiFit system for a period of 3 months. The UbiFit system is shown in Figure 4.

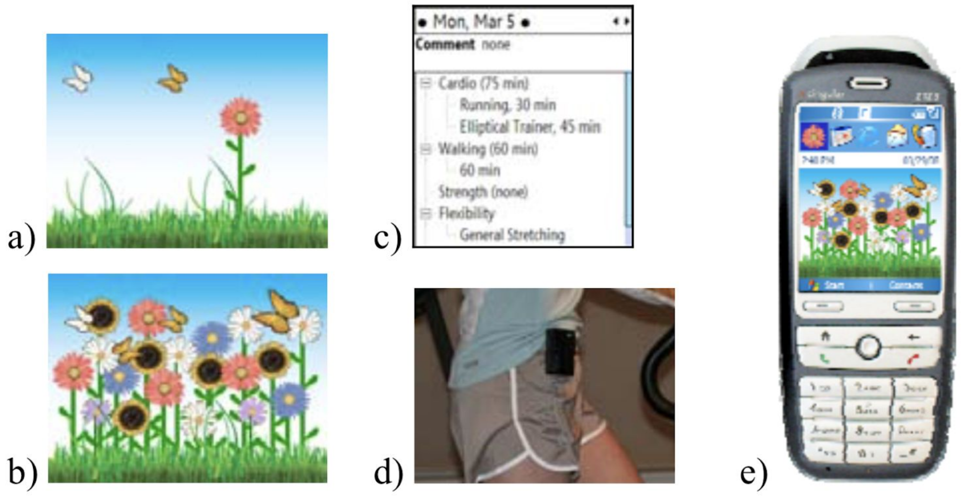


Figure 4. UbiFit system showing glanceable display (a) and (b), mobile application (c) and (e), and activity tracker (d) (Consolvo et al., 2009, p. 1).

The UbiFit system includes a glanceable display, interactive application, and a fitness device to track activity. The glanceable display resides directly in the background screen of the user's mobile phone and uses a stylized "garden" to indicate whether the user has been active or not. Each time the user opens her phone, the garden will be directly visible, and provided that she has been active, the garden will be flourishing. Furthermore, the UbiFit system includes an application, which also resides on the mobile phone and shows more details regarding activity levels and types, and a fitness device which, when worn, automatically tracks activity levels. Following a three month trial period, the researchers interviewed 28 different users and sought insights on goal setting considerations. The strength of Consolvo et al.'s (2009) study stems from participants' actual use of the developed UbiFit system. Insights on persuasive technologies are likely to be much more fruitful when they simulate real world conditions as closely as possible.

Although not directly researching persuasive technology per se Gouveia, Karapanos, and Hassenzahl's (2018) paper on Activity Tracking in Vivo provides very useful insights for the purposes of our study. Gouveia et al. recorded, via wearable cameras, 12 participants' daily use of activity trackers. From this video data, the research team analyzed 244 incidents in which the activity trackers were used, resulting in findings that help explain the behaviors and use cases behind activity trackers. Based on the research, the team suggests the following design considerations: “facilitating learning through glances, providing normative feedback on goal accomplishments, and facilitating micro-plans.” Essentially, most of the time people interact with activity trackers in short ~5 second sessions, they desire for the information to be simple and easy to digest, and they often modify small parts of their behavior throughout the day to accomplish the goals set by the activity tracking software. Although persuasive technology as a concept was not the focus of the study, it is clear to see how the design considerations above can better guide the design of persuasive technologies, especially in the context of wearables.

2.2.4 Ethics in Persuasion

To conclude our review of persuasive technology, we delve into ethical considerations which naturally accompany the creation of persuasive technologies. The ethics of persuasion, when it comes to persuasion via rhetoric, are relatively ubiquitous and understood by most. However, when technology is conducting the persuasion the ethical boundaries are not always as clear.

Authors such as Strauss (1991) have remarked on the ethics of rhetorical persuasion to great extent, stressing the importance of autonomy and deploring non-rational means of persuasion such as false statements of fact. Such ethical aspects of persuasion are not unfamiliar to most, but situations in which technology acts as the persuader and unintended outcomes or consequences may result, the ethical aspects typically require further inspection. For this very

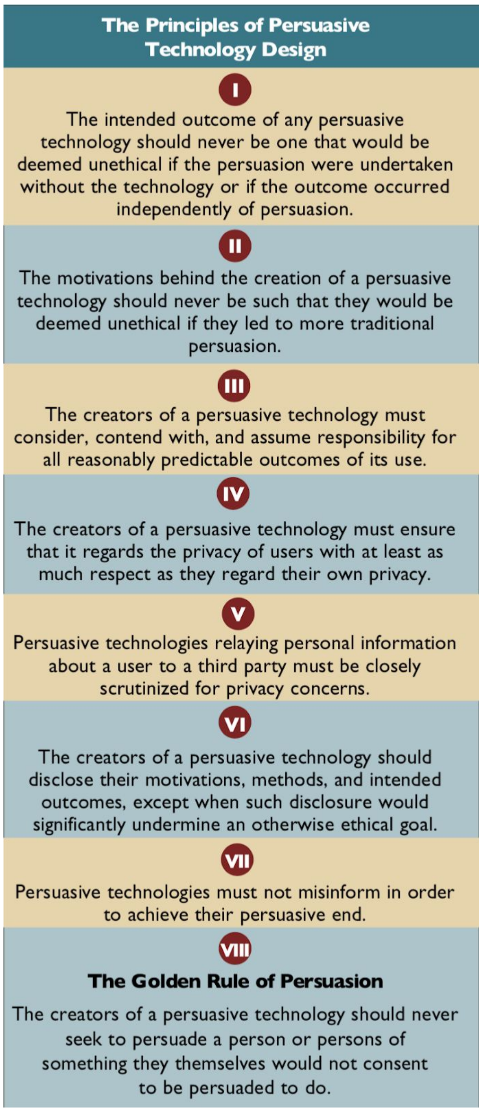


Figure 5. Ethical principles of persuasive design (Berdichevsky & Neuenschwander, 1999, p. 52).

purpose, Berdichevsky and Neuenschwander (1999) conducted a review of the ethics of persuasive technology and drafted a series of principles or guidelines to follow when creating persuasive technology. These principles, shown in Figure 5, shed light onto many of the nuances and unanticipated scenarios that may result from persuasive technology. Notable points of Berdichevsky and Neuenshwander's principles include the designer's responsibility for "all reasonably predictable outcomes" of the persuasive technology's use and the designer's responsibility to safeguard the privacy of the user and the user's personal information.

When creating persuasive technology, it is of utmost importance to be comprehensive in thinking through all possible outcomes of the use of the persuasive technology. How could the technology be misused or misinterpreted? Is there a possibility of forming an addiction or may there be negative consequences to external parties or even the environment? Planning to plead ignorance in the case of an unlikely undesirable outcome is not an effective or ethical strategy.

2.3 Narrative Persuasion

In essence, narrative persuasion can be described as the power stories have to change people's attitudes and beliefs. In the right context, when crafted carefully and deployed deliberately, narratives can be used as a tool to influence and persuade. There is no shortage of research on narratives. In this section, we review how narratives are cognitively processed, which elements are important for narrative persuasion, and how said research is and can be applied to technological contexts.

2.3.1 Cognitive Processing of Narratives

It has been shown that narratives offer "increased comprehension, interest, and engagement" (Dahlstrom, 2014) and that individuals are often more willing to accept normative evaluations from narratives rather than factual arguments (Slater & Rounder, 2002). Narratives are innately persuasive, and because narratives "describe a particular experience rather than general truths, narratives have no need to justify the accuracy of their claims" (Dahlstrom, 2014). As

Dahlstrom further states, “the story itself demonstrates the claim.” This aspect particularly, renders the use of narratives to be a potent, if not dangerous, persuasive design tool.

We can begin to understand why narratives are so effective by breaking down how they are processed in the mind of the reader or observer. Figure 6 illustrates the narrative persuasion process.

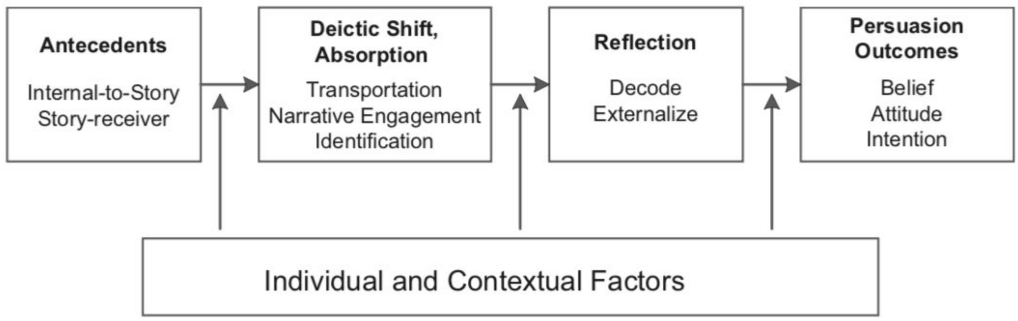


Figure 6. The narrative persuasion process (Hamby, Brinberg, & Jaccard, 2018, p. 114).

Every narrative is affected by the antecedents or prior experiences of the person absorbing the narrative. These antecedents are internal to the story receiver and can be things like education level, gender, or familiarity with the story topic. Once the observer begins to absorb the narrative a deictic shift happens. During this shift the observer is engaged with the story and is “transported” into the story. The observer becomes unaware of herself as an audience member and imagines being one of the characters in the story instead (Cohen, 2001). Following absorption into the story, and typically only once the story has ended, the observer beings to reflect upon it. During this reflection process meaning is constructed from the narrative and the persuasive subtext (Hamby, Brinberg, & Jaccard, 2018). Then in the last stage of successful narrative persuasion, the reflection process yields a change of belief, attitude, or intention that aligns and is consistent with those found within the narrative.

Many models have been proposed to further explain how narratives can influence and persuade, but often narrative persuasion is considered a cloaked form of persuasion, where information is generally accepted first and analyzed in detail only later (Dahlstrom, 2012). Interestingly, another aspect found by Dahlstrom’s research indicates that assertions placed at causal locations within a narrative were accepted as more truthful than the same assertions placed at locations that had no impact on future narrative events (2010).

2.3.2 Narratives in Technology and Health

Narratives embedded within or delivered by technology can serve many functions. They can help explain complicated subjects, increase user engagement or motivation, and even influence users on new beliefs. Particularly in the quantified self movement (where self tracking devices like wearables are used to gather data and draw insights) it has been theorized that narratives could provide a more intuitive form to navigate large amounts of quantitative data (Hilviu & Rapp, 2015). Furthermore, by tapping into episodic memory, data visualized in narrative form can become more memorable (Kwan-Liu Ma et al., 2012). In order to be maximally effective, narratives deployed in a technological context should convey causally connected events, a main animated agent for users to empathize with, and include other elements such as plot twists or contextual details (Hilviu & Rapp, 2015).

Narratives have also been widely used in health related contexts to advocate for healthy behaviors. Although narratives intended to persuade people to stop detrimental behaviors (such as smoking) have not shown significant effects, narratives intended to advocate for detection and prevention behaviors have shown promise (Shen, Sheer, & Li, 2015). Furthermore, when creating narratives for health related contexts, it is advisable to use highly emotional content, as it has been shown to have more effects (Graaf, Sanders, & Hoeken, 2016). Additionally, the content should showcase the healthy behavior as opposed to showing the unhealthy behavior with negative consequences.

2.3.3 Gamification

Although games do not necessarily include narratives, the two are often intertwined and complementary. Gamification in particular has been popularly defined as “the use of game elements to foster behavior change and make mundane tasks more playful (Kappen & Orji, 2018). A variety of these gamification elements or strategies can be used to incentivize a user to reach for instance, health and wellness goals. Shown in Figure 7, Kappen and Orji analyzed five such strategies to be effective for inducing healthy behavior change.

Figure 7. Various gamification strategies used to incentivize users in the process of achieving their health and wellness goals (Kappen & Orji, 2018, p. 53).

Each strategy above has its merits, however, continued usage of such gamification technology may cause an over-dependency and an over-reliance on external motivations (Kappen & Orji, 2018). Alternatively, designers of gamified health technology should opt for designing self-empowerment and self-regulation.

For the purposes of our experimental case study, narrative persuasion holds a lot of promise. Self-awareness is key in helping users change behavior and by interpreting data with visuals or narratives we can help the user do just that (Hilviu & Rapp, 2015). Additionally, gamification elements can be used to further help the user engage with the narrative, leading to deeper immersion and the potential for increased influence and persuasion.

2.4 Measuring UX Behavior

Most new products and attempts at innovation fail due to a lack of upfront work (Cooper, 1999) and a lack of understanding potential users. Not understanding users is especially pertinent for persuasive design where success on the first try is rare and almost always iteration is necessary. Thus, it follows that any intelligence or insight that helps the designer understand the user's behavior will tremendously increase the likelihood of success and increase the speed at which a solution can be found.

There are many methods and metrics, both qualitative and quantitative, that can be used to measure UX behavior. Novel methods utilize technologies such as eye tracking (Bergstrom & Schall, 2014), however most efforts are directed towards measuring engagement. Engagement can be described as the quality of a user experience characterised by heightened attention, positive affect, and intellectual stimulation (O'Brien & Toms, 2008). Interaction of some sort is present, and to some degree, the user is captivated by the technology. As Attfield et al. (2011) state, successful technologies are distinguished by the fact that users are willing to invest time, attention, and emotion into engaging with them.

The key question then becomes, how to design for and measure engagement? Following a review of prior research on engagement, Attfield et al. compiled Table 2, defining eight characteristics of user engagement and appropriate measures for each.

Table 2. Characteristics of user engagement and ways to objectively measure them (Attfield et al., 2011, p. 3).

Characteristic	Definition	Measures
Focused Attention	Focusing attention to the exclusion of other things	Distorted perception of time, follow-on task performance, eye tracking
Positive Affect	Emotions experienced during interaction	Physiological sensors (e.g. face detection)
Aesthetics	Sensory and visual appeal of an interface	Online activity (curiosity-driven behaviour), Physiological sensors (e.g., eye tracking), perceived utility
Endurability	Likelihood of remembering an experience and the willingness to repeat or recommend it	Online activity (e.g. bookmarking, sending emails)
Novelty	Novel, surprising, unfamiliar or unexpected experiences	Physiological sensors (e.g., blood pressure)
Richness and Control	Levels of richness and control	Online activity (e.g., interaction with the site, time spent), Physiological sensors (e.g. mouse pressure)
Reputation, trust and expectation	Global trust users have on a given entity	Online activity (returning user, recommendation)
User Context	User's motivation, incentives, and benefits	Online activity (location, time, past history)

When analyzing UX behavior the feasibility of the method of measurement should be taken into consideration. Face detection and blood pressure monitoring are difficult to implement at scale, however, may yield valuable insights from concentrated focus groups. Online or digital user activity, such as mouse clicks or engagement duration, may not be as robust, but on the other hand can be measured automatically, allowing for data sets that scale exponentially.

2.5 Synthesis

In this section, we reflect on our initial research questions, stated in section 1.2, and attempt to synthesize answers from the literary review conducted above. Following synthesis, we move forward with the development of a hypothesis, which will be inspected in further detail alongside a conceptual framework in Chapter 3.

Persuasive design is the intentional application of psychological principles to induce a desired behavioral change. Ideally persuasive design interventions are conducted ethically and seek to provide benefit not only to the practitioner but to the end user as well. Persuasive design can be used to induce healthy behavior change by increasing the motivation and or the ability of a user to act out a target behavior. If both motivation and ability are present only a trigger is needed. By studying, and measuring if possible, how different interventions are cognitively processed by an end user, a practitioner of persuasive technology can better design for success. Persuasive design or technology can be applied in almost any context, designers however should be wary of selecting target audiences that lack the desire to change or lack experience with the context of the behavioral change intervention in question.

Narratives have been shown to increase comprehension, interest, and engagement (Dahlstrom, 2014), and specifically in the quantified self movement, narratives have been proposed to be effective agents for making data more memorable, more understandable, and more persuasive (Hilviu & Rapp, 2015). The use of narratives in technological contexts have primarily centered around games and gamification. Digital platforms such as mobile devices allow the use of gamification strategies to support and supplement the development of narratives. Quantitative research on the subject, however, is limited. When designing for narrative persuasion, the individual and contextual factors should be taken into consideration. When the narrative is relevant and contextually appropriate for the audience, narrative engagement and transportation is increased, resulting in deeper reflection and a subsequent increase in persuasion outcomes (Hamby, Brinberg, & Jaccard, 2018).

Engagement rate as a mediary step in assessing persuasive design has been called for as early as 2009, but to date, limited quantitative data has been gathered on the topic (Ritterband & Tate, 2009). Most research has focused on theories of behavior change and has largely ignored aspects of user engagement (Short et al., 2015). To fully understand the active mechanisms behind digital persuasive design interventions, both theoretical behavior change and engagement determinants should be studied (Short et al., 2015). To aid in such research Short et al. present a model of user engagement in online interventions shown in Figure 8.

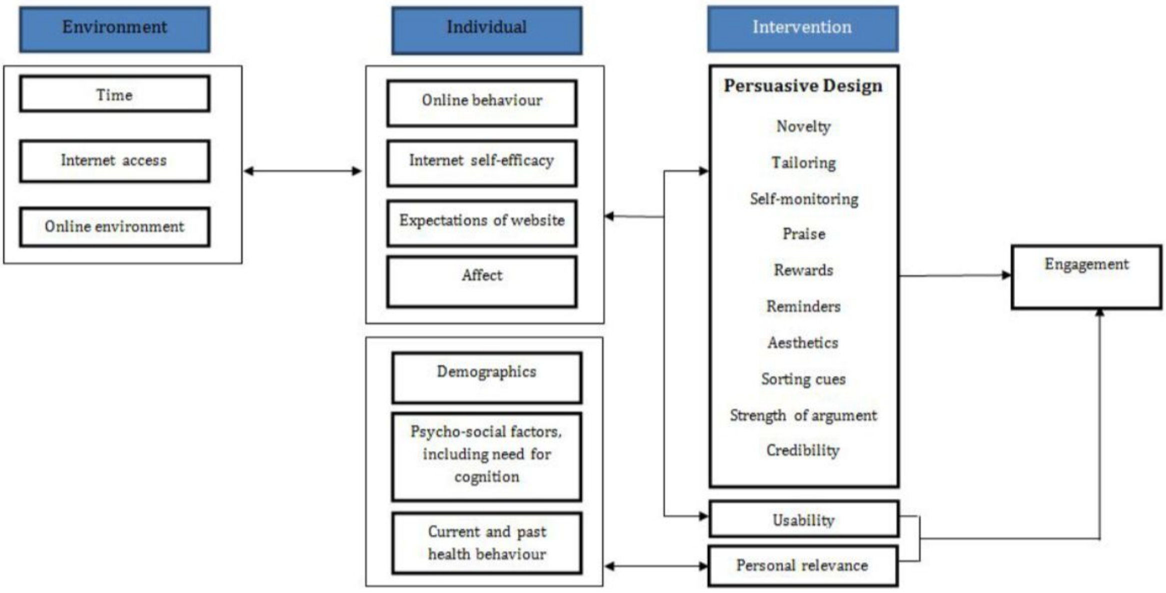


Figure 8. Model of user engagement in online interventions (Short et al., 2015, p. 34).

Short et al. argue that without engagement, all behavior change interventions would fail, so in persuasive design contexts, increased attention should be paid to not only behavior change theory but determinants of engagement as well.

Given the synthesis above, a few themes appear. Psychological theories of behavior change are heavily researched and provide proper groundwork from which to build persuasive design upon. Few studies with quantitative data explaining persuasive design effects exist. Given the feasibility of measuring engagement in digital and technological contexts, engagement has the potential to be a valuable intermediary signal in determining the effectiveness of persuasive design outcomes. How might we apply these themes into the context of the Oura mobile application to induce healthy behavioral change? We hypothesize that by constructing narratives into the application, we can reduce cognitive load and increase engagement, ultimately leading to an increased rate of behavioral change.

Chapter 3

Research Design & Methods

In this chapter, the design of the experimental case study is described, including the methods used to design the case UX as well as collect and analyze data. Firstly, the research questions are reviewed and a conceptual framework is constructed. The research approach is then elaborated upon, using the conceptual framework as a guide. Then the case UX is reviewed followed by the mechanisms used to collect and interpret data. Lastly the research procedure is described.

3.1 Research Questions

Our motives for this research lie in better understanding how persuasive design could be used to help Oura users more easily adopt healthy behaviors. Thus our research questions are formulated accordingly:

> *1) How can persuasive design be used to induce healthy behavioral change?*
>
> *2) Can narrative persuasion specifically, be used to increase engagement in the context of a mobile application?*
>
> *3) Can message timing and format be altered to increase engagement and the rate of ring wear among iOS users of the Oura application?*

Our first question is generic by design and is intended to direct a broad exploration into the topic of persuasive design as it relates to behavioral change, and specifically healthy behavioral change. The literature review conducted in

Chapter 2 presents many perspectives on the topic and sufficiently addresses the question for the purposes of this research.

Our second question dives deeper and focuses on narrative persuasion. Based on the context of this experimental case study, and preliminary literature review, narratives were identified as a potentially high ROI mechanism for delivering persuasion. Our literature review gives a strong starting point from which to answer the question theoretically, quantitative research on the matter however, would provide an even deeper understanding of the phenomena at play.

Our third research question lays the groundwork for quantitative research and further breaks down the mechanisms at play behind behavioral change. Existing academic literature indicates that engagement is indeed a phenomenon worthy of close inspection when designing for persuasion (Short et al., 2015), however the exact relationship between engagement and persuasive design outcomes are not clear, and quantitative research on the subject is warranted.

3.2 Research Design and Approach

In this experimental case study, we follow a deductive research approach. We first examine existing theories on persuasive design, then draft a testable hypothesis, followed by a quantitative experiment to either confirm or reject the hypothesis. The quantitative experiment can best be described as design science, a paradigm in which a design artifact is constructed and applied to obtain knowledge and understanding of a problem domain (Hevner et al., 2004).

For this case, we specifically focus on the onboarding process of the Oura mobile application (further explained in section 3.3), and seek to find out whether constructing the onboarding process in a narrative form will have an affect on user cognition and subsequent user behavior. We hypothesize that an onboarding process presented in a narrative format is easier to digest and absorb, leading to increased transportation and engagement, which then may lead to increased behavioral change. A conceptual framework of the experimental case study is presented in Figure 9.

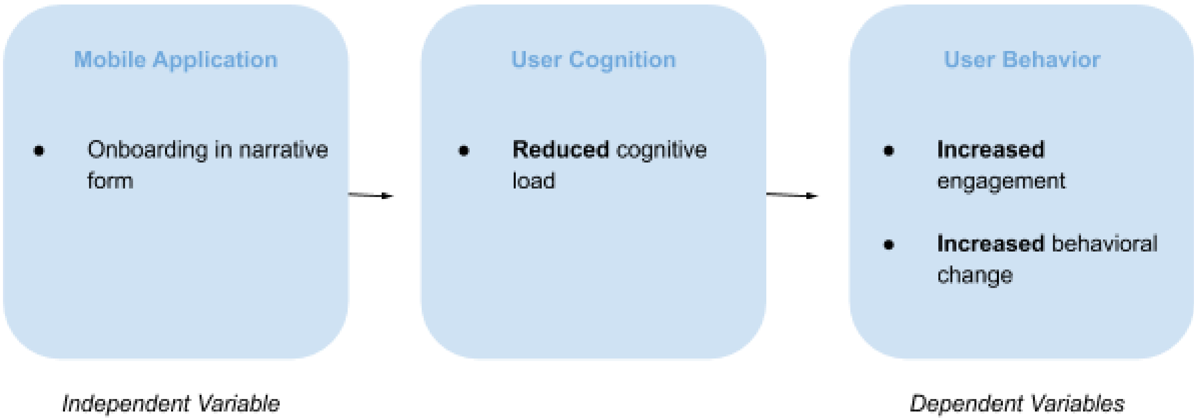


Figure 9. Conceptual framework.

Our independent variable is binary and indicates whether the onboarding process is delivered in a narrative format or a non-narrative format (further explained in section 3.3.1). Our dependent variables are user engagement, as measured by *time in app* and *sessions per day*, and behavioral change, as measured by the *rate of ring wear* (further explained in section 3.4.2).

Our engagement metrics of *time in app* and *sessions per day* are widely used across the app development industry (Mixpanel, 2019). These metrics appropriately quantify and paint a picture of how users on average are interacting with the application. The *rate of ring wear* is a metric specific to Oura and provides an account of how often or not the user actually wears the Oura ring. When creating a behavioral change intervention and selecting a behavior to target, Fogg advises to select the “smallest, simplest behavior that matters” (2009, p. 2). For this research, we identified the *rate of ring wear* to be such a behavior. The *rate of ring wear* is an important metric as wearing the ring is a precursor to the user gaining value from the Oura mobile application. Without wearing the ring and collecting biometric feedback on sleep and recovery, the Oura application is unable to provide the user value-creating insights or guidance.

Thanks to the digital nature of the product at hand, conducting a randomized controlled trial (RCT) is possible. New Oura users are randomly placed into either the treatment or the control group and then proceed to receive the onboarding content in either a narrative or a non-narrative format. Furthermore, users are unaware of participating in a trial of any sort, so the chance of biasing affecting their behavior is slim. This method of testing allows us to test our hypothesis against a null hypothesis while controlling for other variables such as age, gender, timing, or seasonality. In other words, all variables other than the format of the onboarding process are kept the same amongst the treatment and the control group. This allows us to reasonably assume that any effect observed in our engagement or behavioral change metrics is caused by the independent variable.

3.3 Case: UX Design

In this section, we review the nature and design of the case itself. What logic and reasoning stands behind focusing on the onboarding process specifically, and what methods were used to construct the UX of the onboarding process in the treatment and control groups. Arguably, persuasive design is largely delivered through UX design, so close inspection of the UX design process is worthwhile.

Based on the initial qualitative analysis of the Oura application, followed by observations of ring wear and retention rate analytics, the onboarding phase of app use was selected as a frame within which to explore the potential of persuasive design. In terms of creating value for the user, the onboarding process is one of the most, if not the most, important moment in the customer journey. In that moment, first impressions of the product are established, usage habits are formed, and the motivation of the user to learn and understand the product features is the highest. In the case of Oura, as mentioned previously, the ring-wear behavior of users is of high interest (without wearing the ring users can not receive valuable insights). It is presumed that most of the ring-wear habits are formed during the first weeks of app use. From this it follows that by utilizing persuasive design we can nudge users to wear the ring as much as possible in the first weeks of use, leading to sustained ring-wear patterns and habits.

In our experiment, users are randomly assigned to either the treatment or the control group. The control group represents the existing version of the Oura application with a “non-narrative” onboarding process. The treatment group on the other hand, seeks to improve upon the existing onboarding process, by exploring the potential of persuasive design to impact user behavior. The design and methods used to create the treatment group are explained in the following three sections. Procedurally the design process can be broken down into UX strategy, UI design, and copy design. These three phases are ordered chronologically below, however, iteration occured between each phase.

3.3.1 UX Strategy

Before the creation of any design artifact, some planning must occur. Typically, the more thorough the planning and strategic work completed, the better the design is at reaching the desired outcome. In the case of UX design, the process is no different, and UX strategy is often used to describe this pre-work. Specifically, UX strategy takes into consideration user needs, business vision, and technical capabilities to lay out a plan of attack for creating a UX design (Figure 10). This process, as it relates to this case, is described in the following paragraphs.

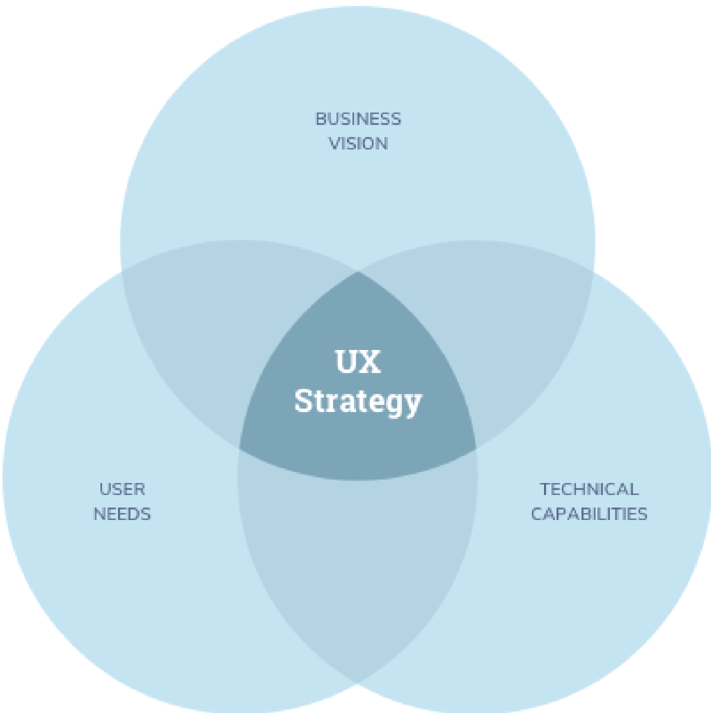


Figure 10. UX Strategy (Tolmie, 2018).

UX strategy begins with analysis of the current situation. We began by heuristically analyzing the Oura mobile application and identifying potential pain points for users. In the existing onboarding process, the following problem areas were identified:

Less than ideal FTUE

The existing first time user experience (FTUE) is underwhelming for the user. The user wonders whether the product is working, and there is no immediate value available for the user. The user is forced to wait for a more valuable interaction at a later time.

High cognitive load

The Oura application is not without its complexities. There are many scientific terms that are likely new for many users and there are many layers of depth in the application to make sense of. During the FTUE, and in the following weeks, all of these create high cognitive load for the user. Reducing the cognitive load would lead to better retention and absorption of the app content and features (Xie et al., 2017).

Timing of user interactions is one-dimensional

Novel and rich content, with which the user can interact, is generated within the app only once per day in the early morning hours. This creates an effective "morning moment" between the user and Oura, but interactions outside of this timeframe feel plain, and leave room for improvements.

Following heuristic analysis of the current situation, we constructed user journey maps, as described by Marquez et al. (2015) to begin visualizing what an ideal onboarding process could look like. The journey maps considered and sought to outline in detail both the users' actions, and the actions performed by Oura (in and outside of the app). Secondly, the user journey maps served to help visualize the touchpoints or interactions between the two, and consider the potential emotions encountered by the user.

Before proceeding to potential solutions from a strategic perspective, technical feasibility was first considered. Technical constraints, namely the availability of software development resources, further guided the development of the UX strategy and the following UI design to narrow in on solutions not requiring complex software development.

Upon completion of the strategic groundwork described above, a few themes surfaced. The need to reduce cognitive load and the need for more timely messaging within the Oura application. From a UX perspective, it became clear that cognitive load could be decreased by breaking existing onboarding messages down into smaller, “bite-sized” pieces. Furthermore, messages could be made more timely by taking into consideration the user’s stage within the user journey, and then crafting and delivering messages accordingly, even during hours outside of the “morning-moment”.

These improvement strategies are well aligned with our intentions of increasing the target behavior of ring wear. Ambient styles of persuasion (requiring little cognitive resources) have been shown to potentially be more effective than more direct forms of persuasion (Ham & Midden, 2010) and just-in-time feedback has been hypothesized to increase health awareness and persuade change in wearable computing applications (Ananthanarayan & Siek, 2012).

3.3.2 UI Design

The process of designing the UI consisted of benchmarking, concepting, and prototyping phases. User testing was limited to informal interviews with friends and colleagues in order to remain agile and keep development speed high. These phases occurred at times concurrently and iteration between each phase took place.

Initially, existing applications both within and outside of the wellness space were reviewed. Interaction mechanics and design language were studied in order to identify successful and unsuccessful components of each. Following benchmarking, a number of different ideas were generated using Sketch (2019), including a variety of card based UIs for breaking onboarding content into “bite-sized” pieces. The best of these ideas were then prototyped using Principal (Principleformac.com, 2019). Following prototyping, technical feasibility was again revisited and the highest ROI design was selected for development.

The chosen design integrates innocuously into the existing design language of the Oura application by delivering content to the home view of the application in a card shaped UI component. Similar to existing UI components this card can be tapped to open more information regarding the topic. The new design, however, breaks longer text-based contents into smaller “swipeable” bites. Additionally, the new design components are delivered into the home view of the application throughout the day, not just once per day in the early morning hours.

3.3.3 Copy Design

A majority of the onboarding content is based on the existing onboarding content, some copywriting was created, however, to help adapt the existing content into the new design. The newly created copy focuses on giving the user context within the application and within the broader onboarding process. As an example, new users are greeted and welcomed to their Oura “journey” and are provided with copy explaining what to expect within the first 24 hours of use. These elements of the copy design not only give context, but also help engage the user into a larger narrative.

Aside from linguistics, timing was an important factor to consider in the copy design. Appropriate timing gives much more weight to the copy and effectively creates the feeling of immersion into a narrative. In designing the timing of the onboarding messages, a few factors were taken into consideration; namely the time of day, stage within the onboarding process, and user specific events such as a low ring battery level. Additionally certain user actions, such as forgetting to wear the ring, the completion of an activity goal, or an exceptionally high score, could trigger a relevant message.

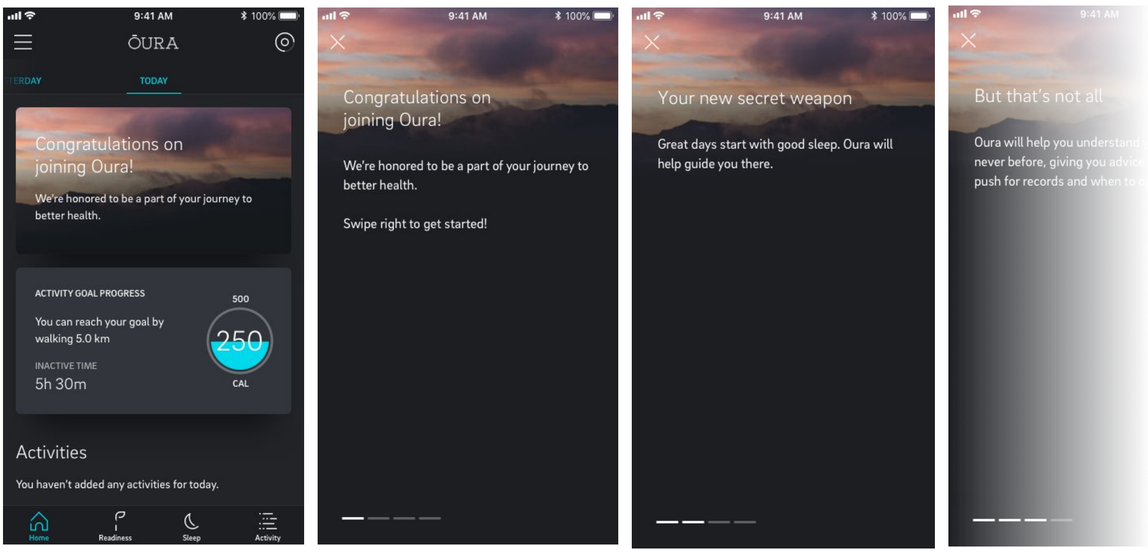
Reflecting once more on the objectives of our research we can see that our design is theoretically well aligned with prior academic literature. With time sensitive messaging, we are able to more intimately consider and design for the needs and emotions of the user, presumably creating more narrative transportation for the user in the process. Secondly, thanks to having more smaller messages throughout the day, we create far more chances for reflection for the user. This narrative transportation and reflection are central in contributing to persuasive design outcomes (Hamby, Brinberg, & Jaccard, 2018, p. 114).

3.4 Data Collection and Analysis

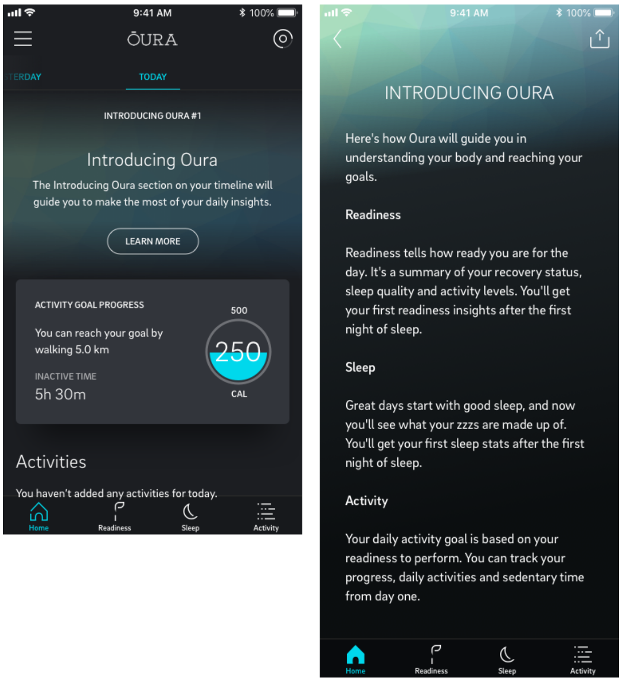
In this section, we discuss how A/B testing was used to collect data. Firstly, the benefits and reasoning behind selecting A/B testing as a data collection method are described, followed by a clarification of the UX variants being tested. Lastly, the test environment is described and the methods used to analyze and process the data are specified.

3.4.1 A/B Testing

As mentioned in section 3.2, the digital nature of the Oura application allowed for a Randomized Controlled Trial (RCT) to feasibly be conducted. When conducted within the IT sphere, RCTs are often referred to as A/B tests, or as Kohavi et al. (2008, p. 3) defines it: “the simplest controlled experiments, where a user is randomly exposed to one of two variants”. The power of an A/B test lies in its ability to determine causality. However, careful consideration must be paid to ensure that consistently the only difference between the two given variants, is the change between the control and the treatment. Differences in treatment (A) and control (B) groups for the experiment were limited to the nature of the presentation of the onboarding content. The visual aspects of these differences are shown in Figure 11.



Treatment (A) - Bite sized content delivered throughout the day



Control (B) - Longer form content delivered once per day

Figure 11. A/B testing narrative vs. non-narrative onboarding process.

In the treatment group (A), bite-sized content explaining product features and user onboarding were delivered to the user throughout the day for a period of 10 days. Alternatively in the control group (B), onboarding content was delivered once per day in the morning for a period of 30 days. Figure 11 shows the visual differences of these variants. The leftmost frame in each variant in Figure 11 displays the home view of the Oura application. In both variants the topmost element in this frame can be tapped to reveal more content. In variant A, this revealed content is succinct and brief, and can furthermore be swiped to reveal additional content. In variant B, all of the content is revealed at once into a scrollable format.

Both variants were randomly assigned to new Oura iOS users, i.e., users that downloaded the Oura application and created an account, for a period of seven days. Users did not know and were not informed of participating in a study of any kind, and it is reasonable to assume that users were not aware of any alternative variation of the onboarding process. One-third of new iOS users were assigned to the treatment (A) group and two-thirds were assigned to the control (B) group. Only iOS users were selected because the variants were only developed for iOS, due to the speed and relative ease of implementation. The percentage of new users assigned to the experimental treatment (A) group was limited to one-third in order to reduce risk and limit any potential negative outcomes caused by the experimental UX.

Group:	**Treatment (A)**	**Control (B)**
% of new users:	33.6%	66.4%
Total # of users:	600	1184

Table 3. Distribution of new Oura iOS users into A/B groups between 08/07/2019-14/07/2019.

3.4.2 Key Performance Indicators

Our KPI's for the experiment are defined in Table 4. These indicators or metrics were measured for both group A and B. User engagement is difficult to quantify and thus both *time in app* and *sessions per day* were recorded. A session is initiated when the Oura app is brought to the foreground (i.e. opened). The session will time out after 30 minutes of inactivity, prompting a new session to be counted if the user opens the app again after the previous session has timed out.

Dependent variables		**Definition**
User engagement:	*Time in app*	Average amount of time (sec) spent in-app per user per day.
	Sessions per day	Average number of sessions per user per day.
Behavioral change:	*Rate of ring wear*	Average percentage of time per user that the ring is worn.

Table 4. Key performance indicators.

3.4.3 Test Environment

The A/B test specifically sought to determine which form of onboarding process is more effective. Since existing users had already experienced the onboarding process, it would have been disruptive to deploy the A/B test to them. Consequently, the deployment of the A/B test was controlled from the server side using Firebase Remote Config, allowing the test to be deployed only to new iOS users, who of course, had not yet proceeded through the app onboarding process

(Firebase, 2019b). Additionally, this aspect of the test environment allowed for the duration of the test recruitment period to be determined on the fly, without the need for new app releases. All in-app analytics events were then recorded to Google Firebase (Firebase, 2019a).

3.4.4 Analysis

Following the completion of data collection, the data was analyzed using SQL queries run on Google's BigQuery tool (Google Cloud, 2019). These queries were used to filter out unfit data points, such as data points stemming from internal testing, and data points generated by users without an actual Oura ring (i.e. users that had downloaded the app but not connected an Oura ring). Additional queries were run to normalize the data. Specifically, the timing/dates associated with each user's data was reordered to ensure that data for each user was collected only for a period of 28 days, starting from the time of ring pairing. For the *rate of ring wear,* data was only observed for a period of 21 days. Finally, SQL queries were used to calculate the desired KPIs from the data set. These queries can be found in Appendix A. Advanced analysis, shown in section 4.3, was conducted with a combination of SQL queries and manipulation in Google Sheets (Google.com, 2019). Users were segmented based off of the number of countries located in during the data collection period. Users located in only one country during the period were labeled as 'domestic' users and users located in two or more countries during the period were labeled as 'international' users. Further analysis of the effect of language on usage data was then conducted by removing international users from the data set. The average *time in app* and *rate of ring wear* for each country was calculated, followed by the calculation of an average for English speaking and non-English speaking countries. A sample of these queries can be found in Appendix B.

Following extraction of the data, the statistical significance of the data was considered. Initially, the standard alpha level of α = 0.05 was chosen. Secondly, the distribution of the population data sets (the control (B) group) were observed for *time in app*, *sessions per day*, and the *rate of ring wear*. Given approximately normal distributions, a Z statistic was then calculated for each KPI using the following equation:

$$z = \frac{\overline{x} - \mu}{\sigma - \sqrt{n}}$$

With Z values in hand, the P values for each KPI were then calculated from a Z score table (Z Score Table, 2019). These P values were then compared with our alpha level of 0.05 to determine statistical significance.

3.5 Research Procedure

Our research process can well be described as a deductive design science. We began by reviewing existing literature on persuasive design and exploring feasible applications in the context of the Oura mobile application. We then hypothesized that by establishing more narrative elements within the onboarding process we could reduce cognitive load for users and engage them more, ultimately leading to an increase in the target behavior of ring wear. Following the formation of our hypothesis, we implemented a design artifact in the form of an altered onboarding process UX. We then deployed the altered UX alongside the existing UX as an A/B test and collected quantitative data on the subsequent user behavior. Lastly, we processed the raw data using standard SQL queries, and then calculated for statistical significance, arriving at our final insights and conclusions.

Chapter 4

Findings

This chapter outlines the quantitative findings from our research. We review the quantitative data from our A/B test, briefly discuss statistical significance, and lastly answer our third and final research question: *can message timing and format be altered to increase engagement and the rate of ring wear among iOS users of the Oura application*?

4.1 Collected Data

After filtering out internal Oura testers and users without an Oura ring the final count of study participants totaled 1,042; 344 users experienced the treatment (A) UX and 698 users experienced the control (B) UX. Over the 28 day data collection period, on average, treatment (A) participants spent 355.6 seconds in the Oura app, an increase of 1.8% compared to the control (B). Additionally, on average, treatment (A) participants engaged in 3.567 sessions per day, an increase of 0.9% over the control (B) group. These metrics indicate that there was an increase in user engagement in the treatment (A) group.

The rate of ring wear in treatment (A) group was 93.89%, this is an increase of 0.8% compared to the control (B) group. In other words, the target behavior of ring wear increased slightly in the treatment (A) group.

4.1.1 Filtered Participants

Group:	**Treatment (A)**	**Control (B)**
% of new users (with ring):	33.0%	67.0%
Total # of user (with ring):	344	698

Table 5. Distribution of users after filtering.

4.1.2 User Engagement

Group:	**Treatment (A)**	**Control (B)**
Time in app (seconds/day):	355.6	349.4

Table 6. Average *time in app* over a 28 day period.

Group:	**Treatment (A)**	**Control (B)**
Sessions per day:	3.567	3.534

Table 7. Average *sessions per day* over a 28 day period.

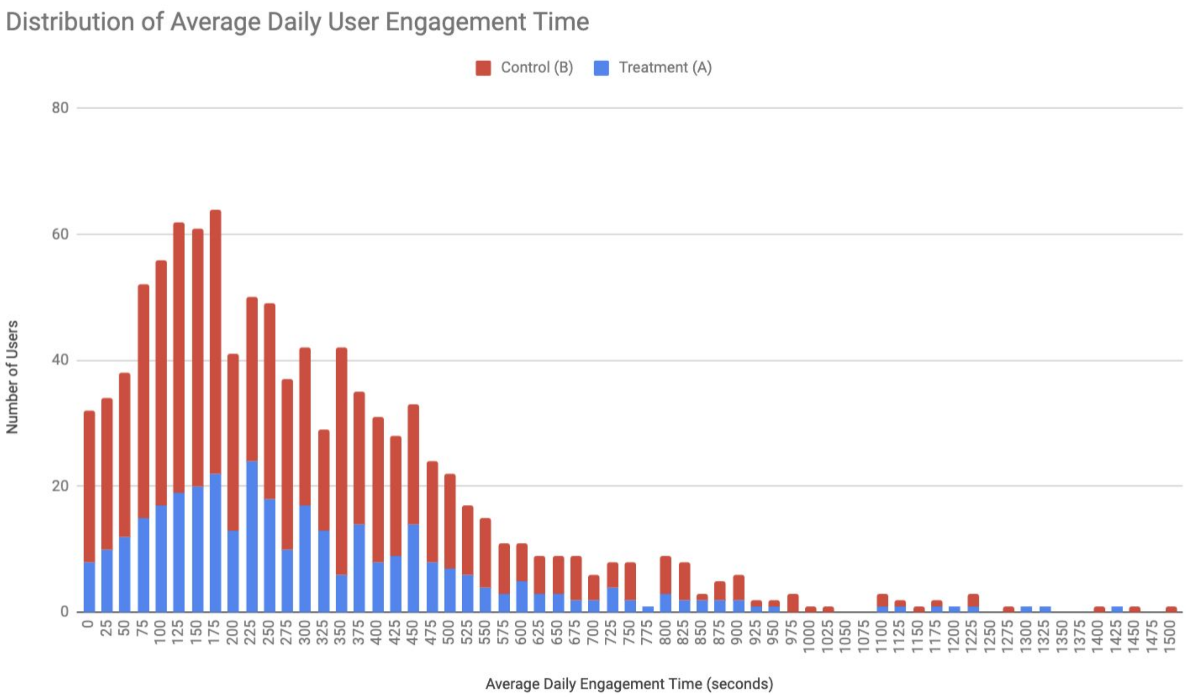
4.1.3 Behavioral Change

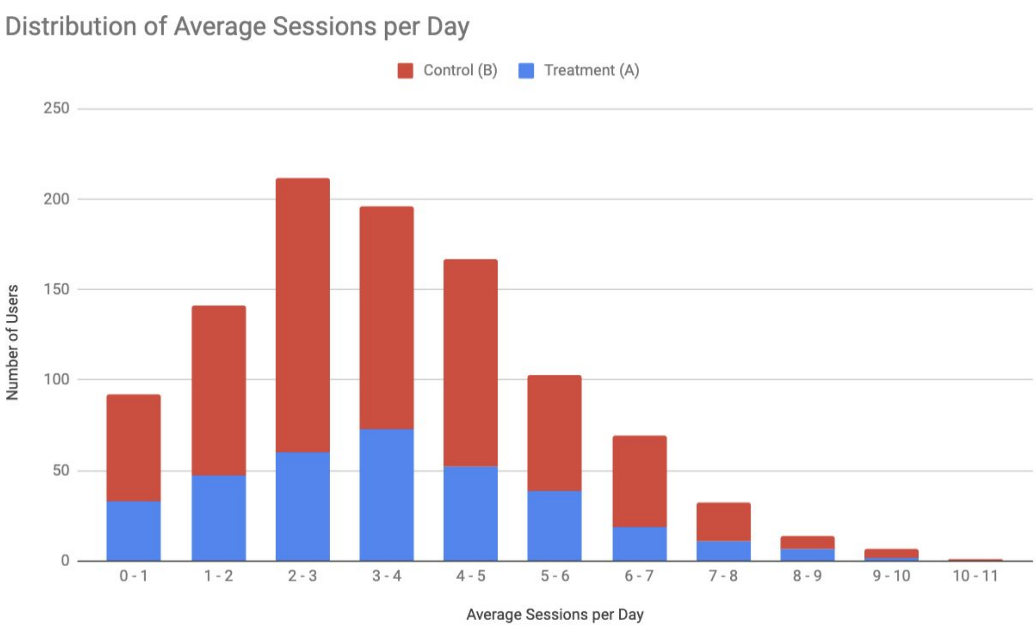
Group:	**Treatment (A)**	**Control (B)**
Rate of ring wear:	93.89%	93.13%

Table 8. Average *rate of ring wear* over a 21 day period (i.e. absolute % of time where ring is worn).

4.2 Statistical Significance

Our null hypothesis states that there is no difference between treatment (A) and control (B) groups. Given our standard alpha level of α = 0.05, in order for us to reject the null hypothesis and determine our results to be statistically significant, the probability of obtaining results as extreme as ours by random chance must be less than 5%. First however, let us observe the distribution of our data sets (Figure 12).





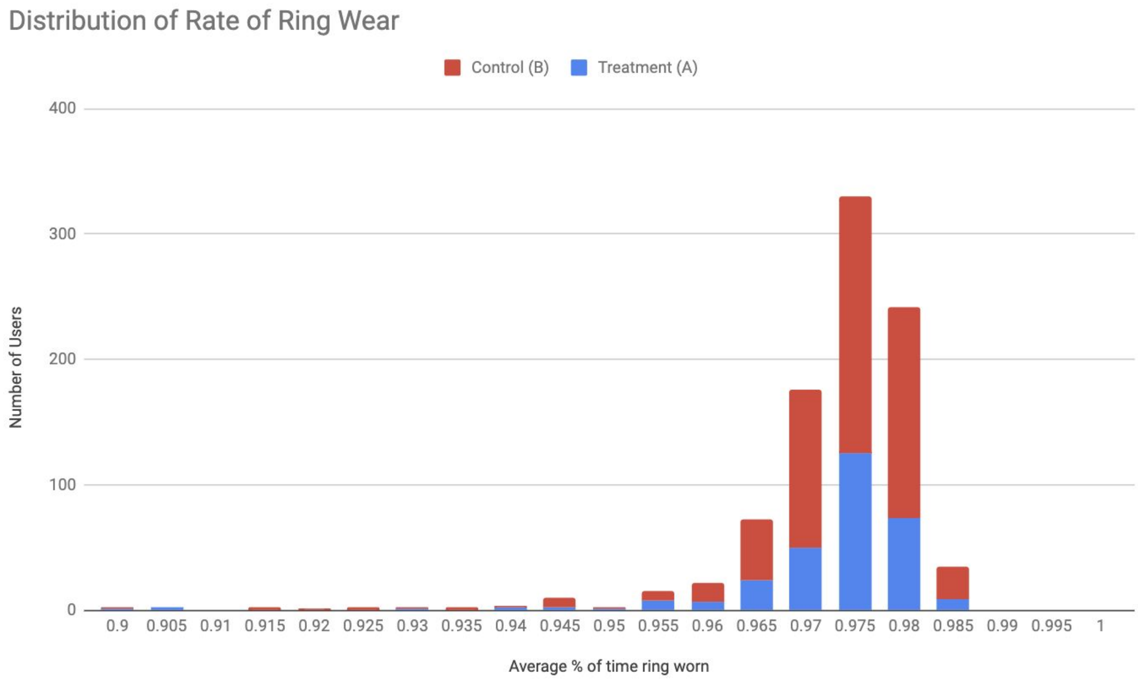


Figure 12. Distribution of *time in app*, *sessions per day*, and *rate of ring wear* data.

From Figure 12 we can see that the distribution of all three KPIs for the control (B) group are approximately normal. Thus, the use of the z-test for normal distribution is valid. The Z test scores for *time in app*, *sessions per day,* and *rate of ring wear* are Z = 0.3289, Z = 0.3133, and Z = 0.8617 respectively. Using a Z table we then find the probability values shown in Table 9.

KPI	P-value	Significance	Statistically significant?
Time in app	0.3745	62.55%	No
Sessions per day	0.3783	62.17%	No
Rate of ring wear	0.1949	80.51%	No

Table 9. Statistical significance for *time in app*, *sessions per day*, and *rate of ring wear*.

As Table 9 shows we are unable to reject the null hypothesis and determine statistical significance. The significance of our treatment for *time in app*, and *sessions per day* is slightly over 60% for each. The significance of our treatment for the *rate of ring wear* is approximately 80%. It is important to note, however, that a lack of statistical significance does not omit the possibility for type II error i.e. we fail to identify a true and real positive effect in the treatment (A) group compared to the control (B) group.

4.3 Advanced Analysis

Advanced analysis of our data reveals several noteworthy aspects regarding engagement and ring wear. Specifically, we find that the effect of our treatment (A) on engagement and ring wear is more profound amongst certain segments of users. These differences are elaborated upon in section 4.3.1 and 4.3.2. First however, we review the breakdown of user session data by device type and country, shown below.

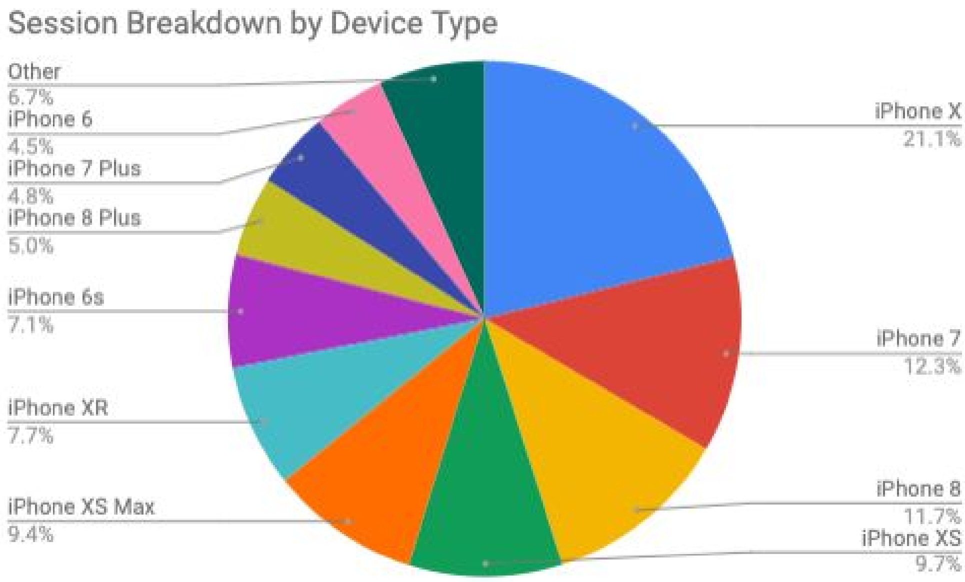


Figure 13. Breakdown of total user sessions by device type.

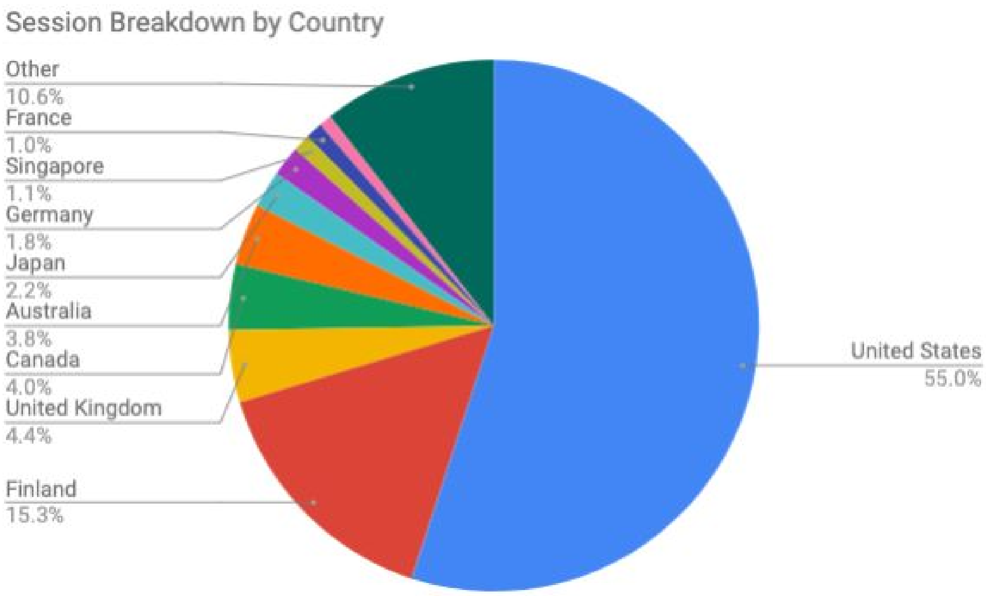


Figure 14. Breakdown of total user sessions by country

Figure 13 and 14 help us understand the context of our data. As designed, all sessions occur on iOS devices. iPhone X, iPhone 7, iPhone 8, and iPhone XS together make up a majority of sessions with older models such as iPhone 6 making up less than 5% of all sessions. A majority of sessions occur in the United States at 55%. Finland makes up a notable share of sessions at 15.3%, followed by countries such as the United Kingdom, Canda, and Australia.

No significant difference in session breakdown by device type exists between treatment (A) and control (B). Things however, become interesting when comparing session breakdown by country between treatment (A) and control (B). These insights are shared in the following section.

4.3.1 Native vs Non-native English Speaking Users

The UIs of both of our experimental groups are relatively text heavy. The treatment (A) group especially leans on an understanding of the English language in order to better construct a narrative and transport the user. By taking the session breakdown by country data, and further classifying by native language, we see an interesting condition emerge. Countries whose native or official language is English, generate a disproportionate amount of sessions in the treatment (A) group compared to control (B).

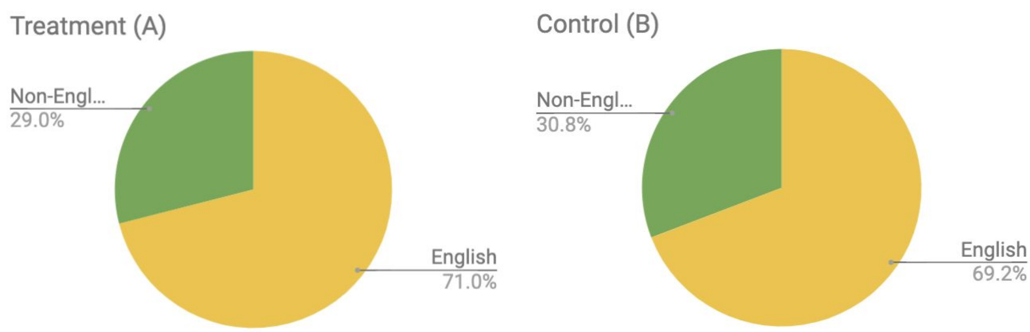


Figure 15. Breakdown of user sessions by English language in treatment (A) and control (B) groups.

It is presumable that users with English language proficiency are more likely to be active and engaged with the language heavy content of the Oura mobile application. Interestingly however, it appears that English speakers are more active, and engaging in more sessions, when experiencing the treatment (A) variant of the application. This implies that the effect of our treatment (A) on the frequency of user sessions, is amplified in the natively English speaking population.

4.3.2 Domestic vs International Users

Taking a closer look at our collected data we partition our users into two groups: domestic users and international users. Domestic users are users with data points from a single country, i.e. users located in only one country during the data collection period. International users are users with data points from multiple countries i.e. users located in two or more countries during the data collection period. We observe that the average *time in app* is slightly higher for international users but especially so for international users in the treatment (A) group (Table 10). Additionally, the *rate of ring wear* is considerably higher for international users (Table 11). International users in the treatment (A) group have the highest *rate of ring wear* at 97.37% ($P < 0.037$) of total time worn.

Group:	**Treatment (A)**	**Control (B)**
Domestic User *Time in app* (seconds/day):	349.8	349.4
International User *Time in app* (seconds/day):	363.4	350.4

Table 10. *Time in app* for domestic and international users.

Group:	**Treatment (A)**	**Control (B)**
Domestic User *Rate of ring wear*:	93.07%	92.90%
International User *Rate of ring wear*:	97.37%	94.17%

Table 11. *Rate of ring wear* for domestic and international users.

Further investigation of Domestic users only:

In order to further explore the impact of geography and/or language on user behavior, we simplify and filter out international users. This reduces the number of participants in the data set from 1,042 to 861. We then look at *time in app* and *rate of ring wear* at the national level, observing the average *time in app* and *rate of ring wear* for each country. Further segmenting by language reveals that *time in app* is considerably higher for English speaking countries, but only when experiencing the treatment (A). The average *rate of ring wear* on the other hand is not increased in English speaking countries. The *rate of ring wear* is, however, elevated in the treatment (A) compared to the control (B) in both segments. Our findings are illustrated in Table 12 and 13.

Country language:	**English**		**Non-English**	
Group:	Treatment (A)	Control (B)	Treatment (A)	Control (B)
Time in app (seconds/day):	490.9	306.5	321.5	311.6

Table 12. Average *time in app* for English and non-English countries.

Country language:	**English**		**Non-English**	
Group:	Treatment (A)	Control (B)	Treatment (A)	Control (B)
Rate of ring wear:	94.26%	90.03%	94.69%	93.24%

Table 13. Average *rate of ring wear* for English and non-English countries.

Compared to control (B), our treatment (A) appears to have a disproportionately larger effect on both *time in app* and *rate of ring wear* in users that are international as opposed to domestic. Additionally, the treatment (A) appears to have a disproportionately larger effect, compared to control (B), on *time in app* amongst English speaking countries. The data above suggests that the treatment (A) is especially effective for users that occasionally travel internationally and are presumably native English speakers. We should, however, exercise caution in interpreting these findings. A considerable portion of the collected data is from users based in the United States. These users may inherently differ considerably from users based in other countries, e.g. on the basis of ad spend or market penetration in that country. These considerations are discussed further in Chapter 5.

4.4 Addressing Research Questions

Given the findings presented in the previous three sections, we can see that the group treated with our persuasive design did in fact engage with the Oura application at a slightly higher rate. This group also slightly increased their rate of ring wear, our target behavior. The effects observed are not large enough to establish statistical significance for the group as a whole. Amongst international users, however, we see that the treatment (A) has an even larger effect on both *sessions per day* and *rate of ring wear.* The lift in *rate of ring wear* amongst international users is statistically significant with $P < 0.037$.

Without full statistical significance we are unable to definitively answer our third research question: *can message timing and format be altered to increase engagement and the rate of ring wear among iOS users of the Oura application?* Although statistical significance was achieved for the *rate of ring wear* among international users, it was not achieved for the broader sample of iOS test subjects in question. Whereas iOS users can reasonably be generalized to Oura users as a whole, international users can not. Qualitatively, however, our findings support the notion that persuasive design, and narrative persuasion in particular, can be used to increase engagement rate in the context of a mobile application.

Chapter 5

Discussion

Ethical persuasive design, in its absolute essence, is more efficient problem solving. It's about guiding and connecting the user to a solution. Helping her solve complex, and sometimes largely behavioral, problems with minimal effort. Persuasive design can help designers, developers, and managers guide users to become more engaged, understand more complex concepts, and ultimately adopt healthier habits.

In our research, we have focused on understanding how persuasive design could be applied in the context of Oura specifically, to induce healthy behavioral change. In the following sections we dive deeper into the theoretical and practical implications of this research, before lastly evaluating the study, and suggesting further areas of interest and future research.

5.1 Theoretical Contributions

Our research began with the simple question of how persuasive design could be used to induce healthy behavioral change. Through our scrutiny of existing academic literature several themes appeared. Narrative persuasion has been extensively practiced and researched in literature and sometimes gaming contexts, but not so extensively in digital contexts relating to product UXs. Secondly, we found that quantitative research exploring how and why a certain persuasive design intervention works is severely lacking. Engagement by many has been called into question as an intermediary step to behavioral change, but research on the topic is limited. These realizations heavily influenced the direction of this research and are mirrored in the theoretical contributions elaborated upon below.

It is unclear whether message timing and format can be altered to increase engagement and the rate of ring wear among iOS users of the Oura application. Statistical significance was not achieved for the effect measured between the treatment (A) and control (B) group. However, when considering only international users, i.e. users located in multiple countries during data collection, a statistically significant ($P < 0.037$) effect on the *rate of ring wear* is observed. Further observing geographic data implies that the effect of the treatment (A) is accentuated in English speaking populations. Based on the subjective analysis of these results, theoretical contributions are discussed in the following paragraphs.

Seemingly small variances in the UX of a digital product can have disproportionate effects on the way users engage and interact with a product. The pareto principle applies here as well. Big design interventions are not always needed to affect user behavior. Although statistical significance was not achieved, our research suggests that relatively simple design changes, such as the timing of messages, have the potential to drive authentic user engagement and subsequently impact a target user behavior. Message timing within a digital UX can be extremely effective as it dictates the rhythm and pattern of interaction between the user and the product. Instead of forcing the user to bend to an arbitrary schedule defined by the technology, with careful consideration, the technology can be designed to fit into the rhythm and schedule of the user's life.

Capturing and understanding variations, if any exist, in how different user segments respond to a persuasive design intervention, can be a beacon for further product development. Certain segments of users may respond more favorably and others more negatively. Access to granular usage data allows for a deeper understanding of user interactions and can guide future product development. In our research, we observed *time in app* and the *rate of ring wear* increase by 1.8% ($P < 0.375$) and 0.8% ($P < 0.195$) respectively, in the treatment (A) group. When considering only international users, *time in app* and the *rate of ring wear* increased by 3.7% ($P < 0.378$) and 3.4% ($P < 0.037$) respectively, in the treatment (A) group. This insight indicates that the treatment (A) was more effective amongst international users. Furthermore, it implies that the engagement metric of *time in app* more effectively serves as a proxy for future behavioral change in international users.

Measures of engagement show potential in being used as leading indicators for the effectiveness of a persuasive design intervention. In the context of digital products especially, various engagement metrics should be captured and analyzed to assist in the development of a persuasive design. Our research supports the notion that an increased engagement rate might play a role in the subsequent facilitation of behavioral change. In our research the treatment (A) group showed higher average engagement, as measured by *time in app* and *sessions per day,* although not in a statistically significant manner. *Time in app* and *sessions per day* were observed to increase by 1.8% ($P < 0.375$) and 0.9% ($P < 0.378$) respectively. Additionally, the treatment group showed an increase in the *rate of ring wear* by 0.8% ($P < 0.195$). The mediary effect of increased engagement was even larger among international users as mentioned in the previous paragraph. Mediary engagement metrics should however, be chosen with caution. They should only serve to help understand and make sense of the interaction between the user and the product. In the case of some products, it very well may be reasonable for a certain engagement metric to decrease in order for a subsequent behavior metric to increase.

5.2 Practical Implications

The magic of persuasive design lies in its ability to assist the user seamlessly and effortlessly, typically flying far below the radar of conscious thought. This adds to the allure of persuasive design. We should however, be cautious of thinking of persuasive design as a 'magic bullet' for a product in despair. In this section, we assemble the practical implications of our research. We begin by assessing concrete findings from the research, and end with more abstract generalizations.

Despite a null result in the broader sample of participating iOS users, altering message timing and format to better engage and guide users to healthy behaviors, shows promise. International users responded particularly well to the treatment (A) and showed a 96.33% significant increase in the *rate of ring wear.* Clearly then, the modification of messaging attributes such as timing or format, have the power to distinctively impact the relationship between app and user. Further research and iteration can bear fruits in designing a UX that is positively more engaging and more effectively guides users to healthy habits and behaviors.

In aggregate, international users may be better aligned with an 'early adopter' or early 'majority mindset', leading to higher rates of engagement and processing of onboarding content, ultimately resulting in higher rates of ring wear. Whereas the domestic users experience no change in *time in app* between treatment (A) and control (B), international users experience an increase of 3.7% ($P < 0.378$) (Table 10). Practically no change in *time in app* is seen between international and domestic users in the control (B) group leaving us to believe that the increase in engagement is caused by the timing and format of the onboarding messaging. Given this finding, we propose that if a design intervention significantly improves the UX for an international user, without detracting from the UX of a domestic user, this design should be implemented.

Our study finds no result among the general sample studied and a marginally statistically significant result for the subset of international users in the sample studied. On this basis, the treatment (A) variation of the Oura application should not be implemented to the broader population of both iOS and Android users, but should instead be thought of us a starting point for further iteration. A larger, statistically significant, effect should be observed before implementing. Furthermore, given that domestic users account for a majority of our users studied at 82%, a design intervention should at least create an incremental improvement in the UX for this majority, in order to be worthy of implementation.

In conclusion, persuasive design can be thought of as a framework from which to approach the design of products and services. In and of itself persuasive design will not resurrect a product or feature that itself is not intrinsically valuable to the user. In the context of Oura, persuasive design can be used to actively engage the user, and increase micro behaviors, such as the rate of wear of an Oura ring. In the future, these micro behaviors can be leveraged to help induce larger keystone behaviors, such as going to bed earlier, to ultimately craft lasting healthy habits for the user. Designing for large or drastic behavioral changes can be incredibly challenging. By deliberately designing for intermediary behavioral changes, the wheels of progress can be set in motion and learning can begin.

5.3 Evaluation of the Study

As with every study, this study is not without its limitations. In this section we evaluate the design, approach, and methods used in the study. The dangers of over generalizing data, the implications of partial statistical significance, and considerations for the improvement of this study are all discussed.

Our research began with the motivation to better understand how persuasive design could be used to help Oura users more easily adopt healthy behaviors. Furthermore, we sought to productize findings from theoretical research into the Oura mobile application and quantitatively measure any changes in user behavior. In that regard, this research has served us well, as we have managed to quantitatively experiment with real Oura ring users of the Oura mobile application. The data collected from that experiment has allowed us to investigate the interplay between narratives, user engagement, and subsequent behavioral change, namely, the rate at which a user chooses to wear their Oura ring. Although our persuasive design variant (A) did not yield statistically significant improvement over control (B) this time, it has afforded us a much deeper understanding of how persuasive design, and narratives in particular, could be used to help Oura ring users adopt healthier habits.

The methods used to collect and analyze our data are quite steadfast. Google Firebase is an industry standard for collecting in app usage data and BigQuery is commonly used to post process this data. SQL is a widely used programming language for managing databases and large data sets. In other words the data generated from this research is reliable. Data itself however, can be inherently difficult to analyze and interpret. The risks of data misinterpretation in this study our outlined in the following paragraphs.

First and foremost we must note that iOS users, although they account for the majority, may not fully represent the whole of Oura mobile application users. On average, iOS users have more disposable income and are more skewed to geographic locations in the western world than Android users. It is difficult to say how different, if at all, our results would have been if our study participants were Android users.

Secondly, we draw into attention the impact that marketing and ad spend may have on user behavior, as it relates to geography. Heavy marketing and ad spend in a certain geography may attract different types of users in that geography compared with users from another geography that gain awareness of the product through other means such as personal research. It is likely that market penetration in countries such as the U.S. is at a much later stage than countries such as Australia, or Switzerland. This reflects in the user behavior of participants from the U.S. (*time in app*: 331.4 sec/day) compared to say that of Australia (*time in app*: 408.2 sec/day). Users from obscure markets are likely to be much more engaged and proactive compared with users from markets where the product has reached a level of pervasiveness. This characteristic of the data is important to keep in mind, especially when considering the fact that 55% of user sessions in the study were conducted in the United States (Figure 14).

Statistical significance was not achieved for our key performance indicators of *time in app*, *sessions per day,* or *rate of ring wear* for our sample population as a whole. On the other hand, the *rate of ring wear* was found to increase statistically significantly for international users. We can not therefore rely on these results as being definitive. Given a significant result within a subset of users, however, we can safely assume that there is considerable potential in the power of narrative persuasion, to induce healthy behavioral change among Oura users.

In a replication of this study, from an academic point of view, it would be advised to collect data from a larger segment of the whole user base, or to collect data over a longer period of time in order to achieve statistical significance. From a practical point of view however, it would be advised to design more daring and risky persuasive design interventions, that clearly and significantly differ from the control. This type of approach is more likely to yield significant results one way or another. On one end, a clear improvement is found and on the other, a clear lesson is learned.

5.4 Further Research

In reality, persuasive design is an iterative process. A persuasive design is hardly ever successful on the first try, and variations of a design are typically needed to yield even marginal shifts in user behavior. Our research provides a solid base from which to start further exploring the role that narratives play, and the impact that narrative persuasion has, on user behavior in the context of digital products. Interesting avenues of further research are delineated below.

We found a clear difference in domestic and international users of the Oura app, our segmentation of these users by single or multiple countries visited however, was arbitrary. Exploring this relationship further, by studying specific country combinations that yield significantly different results, could be a worthy pursuit. This may also give insight into better understanding why international users are more engaged than domestic users.

Our simple analysis of the collected data on a country level, revealed a difference in engagement between English and non-English speaking countries. What role exactly does language play in this equation? Could in app texts be localized to the native languages of each user? How might changes such as these affect engagement, and will the associated costs be justified?

The world we live in is becoming ever more saturated with stimulus. In 1984 Cialdini stated that we will have to depend increasingly on our shortcuts to navigate the overstimulated world. How could we better utilize narratives to form more efficient ‘shortcuts’ to reduce stimulus and reduce cognitive load? Message timing and format are but a few of the attributes of narratives that could be studied. Narratives and stories are inherent to being human. They hold great potential and can play a key role in helping people become healthier. We can continue making technology easier, more effective, and more enjoyable to use, simply by humanizing it through stories and narratives.

Bibliography

Ajzen, I. and Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84(5), pp.888-918.

Ananthanarayan, S. and Siek, K. (2012). Persuasive Wearable Technology Design for Health and Wellness. In: *6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*.

Aristotle (350 B.C.E). *The Internet Classics Archive | Rhetoric by Aristotle*. [online] Classics.mit.edu. Available at: http://classics.mit.edu/Aristotle/rhetoric.1.i.html [Accessed 2 Apr. 2019].

Attfield, S., Kazai, G., Lalmas, M. and Piwowarski, B. (2011). Towards a science of user engagement. In: *WSDM'11*.

Berdichevsky, D. and Neuenschwander, E. (1999). Toward an ethics of persuasive technology. *Communications of the ACM*, 42(5), pp.51-58.

Bergstrom, J. and Schall, A. (2014). *Eye tracking in user experience design*. Waltham, MA: Elsevier.

Berkovsky, S., Freyne, J. and Oinas-Kukkonen, H. (2012). Influencing Individually: Fusing Personalization and Persuasion. *ACM Transactions on Interactive Intelligent Systems*, 2(2), pp.4153-4157.

Cialdini, R. (1984). *Influence, The Psychology of Persuasion*. New York: Harper Collins.

Cohen, J. (2001). Defining Identification: A Theoretical Look at the Identification of Audiences With Media Characters. *Mass Communication and Society*, 4(3), pp.245-264.

Consolvo, S., Klasnja, P., McDonald, D. and Landay, J. (2009). Goal-Setting Considerations for Persuasive Technologies that Encourage Physical Activity. In: *Persuasive*.

Cooper, R. (1999). The Invisible Success Factors in Product Innovation. *Journal of Product Innovation Management*, 16(2), pp.115-133.

Dahlstrom, M. (2012). The Persuasive Influence of Narrative Causality: Psychological Mechanism, Strength in Overcoming Resistance, and Persistence Over Time. *Media Psychology*, 15(3), pp.303-326.

Dahlstrom, M. (2010). The Role of Causality in Information Acceptance in Narratives: An Example From Science Communication. *Communication Research*, 37(6), pp.857-875.

Dahlstrom, M. (2014). Using narratives and storytelling to communicate science with nonexpert audiences. *Proceedings of the National Academy of Sciences*, 111(Supplement_4), pp.13614-13620.

Fazio, R., Zanna, M. and Cooper, J. (1978). Direct Experience and Attitude-Behavior Consistency: An Information Processing Analysis. *Personality and Social Psychology Bulletin*, 4(1), pp.48-51.

Firebase. (2019a). *Firebase*. [online] Available at: https://firebase.google.com/ [Accessed 24 Jun. 2019].

Firebase. (2019b). *Firebase Remote Config | Firebase*. [online] Available at: https://firebase.google.com/docs/remote-config [Accessed 24 Jun. 2019].

Fogg, B. (2009a). A Behavior Model for Persuasive Design. In: *Proceedings of the 4th International Conference on Persuasive Technology*.

Fogg, B. (2009b). Creating Persuasive Technologies: An Eight-Step Design Process. In: *Persuasive '09*.

Fogg, B. (2003). *Persuasive Technology - using Computers to Change What We Think and Do*. San Francisco: Morgan Kaufmann Publishers.

Geddes, J. (2016). *Elaboration Likelihood Model Theory – Using ELM to Get inside the User's Mind*. [online] The Interaction Design Foundation. Available at: https://www.interaction-design.org/literature/article/elaboration-likelihood-model-theory-using-elm-to-get-inside-the-user-s-mind [Accessed 3 Apr. 2019].

Google Cloud. (2019). *BigQuery - Analytics Data Warehouse | BigQuery | Google Cloud*. [online] Available at: https://cloud.google.com/bigquery/ [Accessed 24 Jun. 2019].

Google.com. (2019). *Google Sheets: Free Online Spreadsheets for Personal Use*. [online] Available at: https://www.google.com/sheets/about/ [Accessed 27 Sep. 2019].

Gouveia, R., Karapanos, E. and Hassenzahl, M. (2018). Activity Tracking in vivo. Montreal, QC, Canada: Association for Computing Machinery.

Graaf, A., Sanders, J. and Hoeken, H. (2016). Characteristics of Narrative Interventions and Health Effects: A Review of the Content, Form, and Context of Narratives in Health-related Narrative Persuasion Research. *Review of Communications Research*, 4, pp.89-130.

Halko, S. and Kientz, J. (2010). Personality & Persuasive Technology: An Exploratory Study on Health-Promoting Mobile Applications. In: *Persuasive*. Seattle, Washington.

Hamby, A., Brinberg, D. and Jaccard, J. (2018). A Conceptual Framework of Narrative Persuasion. *Journal of Media Psychology*, 30(3), pp.113-124.

Ham, J. and Midden, C. (2010). Ambient Persuasive Technology Needs Little Cognitive Effort: The Differential Effects of Cognitive Load on Lighting Feedback versus Factual Feedback. In: *Persuasive*. Eindhoven, The Netherlands: Springer-Verlag Berlin Heidelberg.

Hevner, A., March, S., Park, J. and Ram, S. (2004). Design Science In Information Systems Research. *MIS Quarterly*, 28(1), pp.75-105.

Hilviu, D. and Rapp, A. (2015). Narrating the quantified self. In: *UbiComp/ISWC '15 Adjunct*.

Kappen, D. and Orji, R. (2017). Gamified and persuasive systems as behavior change agents for health and wellness. *XRDS: Crossroads, The ACM Magazine for Students*, 24(1), pp.52-55.

Kohavi, R., Longbotham, R., Sommerfield, D. and Henne, R. (2008). Controlled experiments on the web: survey and practical guide. *Data Mining and Knowledge Discovery*, 18(1), pp.140-181.

Kwan-Liu Ma, Liao, I., Frazier, J., Hauser, H. and Kostis, H. (2012). Scientific Storytelling Using Visualization. *IEEE Computer Graphics and Applications*, 32(1), pp.12-19.

Latham, G. and Locke, E. (1991). Self-regulation through goal setting. Organizational Behavior and Human Decision Processes, 50(2), pp.212-247.

Marquez, J., Downey, A. and Clement, R. (2015). Walking a Mile in the User's Shoes: Customer Journey Mapping as a Method to Understanding the User Experience. *Internet Reference Services Quarterly*, 20(3-4), pp.135-150.

Mixpanel. (2019). 9 Important User Engagement Metrics to Measure for Apps | Topics | Mixpanel. [online] Available at: https://mixpanel.com/topics/important-user-engagement-metrics-apps/ [Accessed 29 May 2019].

O'Brien, H. and Toms, E. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6), pp.938-955.

Oinas-Kukkonen, H. and Harjumaa, M. (2009). Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems, 24.

Oinas-Kukkonen, H. and Harjumaa, M. (2008). Towards Deeper Understanding of Persuasion in Software and Information Systems. In: *First International Conference on Advances in Computer-Human Interaction*.

Petty, R. and Cacioppo, J. (1986). The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology*. Cambridge, Massachusetts: Academic Press

Principleformac.com. (2019). [online] Available at: https://principleformac.com/ [Accessed 24 Jun. 2019].

Proust, M., Mayor, A., Kilmartin, T., Enright, D. and Kilmartin, J. (2000). In search of lost time. London: Vintage.

Ritterband, L. M., & Tate, D. F. (2009). The science of internet interventions. *Annals of Behavioral Medicine*, 38(1), 1-3.

Shen, F., Sheer, V. and Li, R. (2015). Impact of Narratives on Persuasion in Health Communication: A Meta-Analysis. *Journal of Advertising*, 44(2), pp.105-113.

Short, C., Rebar, A., Plotnikoff, R. and Vandelanotte, C. (2015). Designing engaging online behaviour change interventions: A proposed model of user engagement. *The European Health Psychologist*, 17(1), pp.32-38.

Sketch. (2019). *The digital design toolkit*. [online] Available at: https://www.sketch.com/ [Accessed 24 Jun. 2019].

Slater, M. and Rouner, D. (2002). Entertainment Education and Elaboration Likelihood: Understanding the Processing of Narrative Persuasion. *Communication Theory*, 12(2), pp.173-191.

Strauss, D. (1991). Persuasion, Autonomy, and Freedom of Expression. *Columbia Law Review*, 91(2), p.334-371.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. Cognitive Science, 12(2), pp.257-285.

Sweller, J., Merrienboer, J. and Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), pp.251-296.

Tolmie, K. (2018). *What is User Experience Strategy*. [online] Medium. Available at: https://medium.com/@MentorMate/what-is-user-experience-strategy-7af34d694fbb [Accessed 10 Jun. 2019].

Torning, K. and Oinas-Kukkonen, H. (2009). Persuasive System Design: State of the Art and Future Directions.

Xie, H., Wang, F., Hao, Y., Chen, J., An, J., Wang, Y. and Liu, H. (2017). The more total cognitive load is reduced by cues, the better retention and transfer of multimedia learning: A meta-analysis and two meta-regression analyses. *PLOS ONE*, 12(8), p.e0183884.

Z Score Table. (2019). Z Table and Z score calculation. [online] Available at: http://www.z-table.com/ [Accessed 7 Sep. 2019].

Appendix A

SQL queries for collected data.

TIme in app

Generating onboarding_study_participant_data table

LINK SHARING HIDE EDITOR FULL SCREEN

```sql
SELECT *
FROM `oura-69259.analytics_164364211.events_*`, UNNEST(user_properties) as up

WHERE _TABLE_SUFFIX BETWEEN '20190708' AND '20190812'
AND up.key = 'onboarding_study_group'

AND user_pseudo_id IN (SELECT DISTINCT user_pseudo_id FROM `oura-69259.analytics_164364211.onboarding_study_participant_data`, UNNEST(event_params) as
ep WHERE event_name = 'screen_view' AND ep.value.string_value = 'onboarding_ring_found')
```

Filtering out unreal users

LINK SHARING HIDE EDITOR FULL SCREEN

```sql
DELETE
FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data`

WHERE user_pseudo_id IN
(SELECT DISTINCT user_pseudo_id FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data`, UNNEST(event_params) as ep WHERE event_name
= 'screen_view' AND ep.value.string_value = 'onboarding_ring_found' AND ep.key = 'firebase_screen' AND event_timestamp > 1563148799000000)

OR user_pseudo_id IN (SELECT user_pseudo_id FROM (SELECT user_pseudo_id, SUM(engagement_time)/1000 as total_engagement_time FROM `oura-
69259.analytics_164364211.v2_onboarding_study_user_engagement_data` GROUP BY user_pseudo_id) WHERE total_engagement_time < 200)
```

Generating onboarding_study_user_engagement_data

LINK SHARING HIDE EDITOR FULL SCREEN

```sql
SELECT DISTINCT o.user_pseudo_id, o.value.string_value as onboarding_study_group, o.event_date, o.event_name, (SELECT value.int_value FROM
UNNEST(event_params) WHERE key = "engagement_time_msec") AS engagement_time, f.first_real_time_stamp, o.event_timestamp, ((o.event_timestamp -
f.first_real_time_stamp)/1000000)/86400 as day_number

FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data` as o, (SELECT user_pseudo_id, MIN(event_timestamp) as first_real_time_stamp
  FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data`, UNNEST(event_params) as ep WHERE event_name = 'screen_view' AND
ep.value.string_value = 'onboarding_ring_found' AND ep.key = 'firebase_screen' GROUP BY user_pseudo_id) as f

WHERE f.user_pseudo_id = o.user_pseudo_id
AND o.event_name = 'user_engagement'
AND ((o.event_timestamp - f.first_real_time_stamp)/1000000)/86400 > 0
AND ((o.event_timestamp - f.first_real_time_stamp)/1000000)/86400 < 28

ORDER BY user_pseudo_id, event_timestamp
```

engagement time

LINK SHARING HIDE EDITOR FULL SCREEN

```sql
SELECT onboarding_study_group, (AVG(total_engagement_time)/1000)/28 as engagement_time_daily
FROM
(SELECT user_pseudo_id, onboarding_study_group, SUM(engagement_time) as total_engagement_time

FROM `oura-69259.analytics_164364211.v4_onboarding_study_user_engagement_data`

GROUP BY user_pseudo_id, onboarding_study_group
)
GROUP BY onboarding_study_group
ORDER BY onboarding_study_group
```

Sessions per day

Generating onboarding_study_participant_data table

LINK SHARING HIDE EDITOR FULL SCREEN

```sql
SELECT *
FROM `oura-69259.analytics_164364211.events_*`, UNNEST(user_properties) as up

WHERE _TABLE_SUFFIX BETWEEN '20190708' AND '20190812'
AND up.key = 'onboarding_study_group'

AND user_pseudo_id IN (SELECT DISTINCT user_pseudo_id FROM `oura-69259.analytics_164364211.onboarding_study_participant_data`, UNNEST(event_params) as
ep WHERE event_name = 'screen_view' AND ep.value.string_value = 'onboarding_ring_found')
```

Filtering out unreal users   LINK SHARING   HIDE EDITOR   FULL SCREEN

```
DELETE
FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data`

WHERE user_pseudo_id IN
(SELECT DISTINCT user_pseudo_id FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data`, UNNEST(event_params) as ep WHERE event_name
= 'screen_view' AND ep.value.string_value = 'onboarding_ring_found' AND ep.key = 'firebase_screen' AND event_timestamp > 1563148799000000)

OR user_pseudo_id IN (SELECT user_pseudo_id FROM (SELECT user_pseudo_id, SUM(engagement_time)/1000 as total_engagement_time FROM `oura-
69259.analytics_164364211.v2_onboarding_study_user_engagement_data` GROUP BY user_pseudo_id) WHERE total_engagement_time < 200)
```

Generating onboarding_study_session_start_data table   LINK SHARING   HIDE EDITOR   FULL SCREEN

```
SELECT DISTINCT o.user_pseudo_id, o.value.string_value as onboarding_study_group, o.event_date, o.event_name, f.first_real_time_stamp,
o.event_timestamp, ((o.event_timestamp - f.first_real_time_stamp)/1000000)/86400 as day_number

FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data` as o, (SELECT user_pseudo_id, MIN(event_timestamp) as first_real_time_stamp
  FROM `oura-69259.analytics_164364211.v2_onboarding_study_participant_data`, UNNEST(event_params) as ep WHERE event_name = 'screen_view' AND
ep.value.string_value = 'onboarding_ring_found' AND ep.key = 'firebase_screen' GROUP BY user_pseudo_id) as f

WHERE f.user_pseudo_id = o.user_pseudo_id
AND o.event_name = 'session_start'
AND ((o.event_timestamp - f.first_real_time_stamp)/1000000)/86400 > 0
AND ((o.event_timestamp - f.first_real_time_stamp)/1000000)/86400 < 28

ORDER BY user_pseudo_id, event_timestamp
```

average sessions/day   LINK SHARING   HIDE EDITOR   FULL SCREEN

```
SELECT onboarding_study_group, AVG(sessions_per_day) as average_sessions_per_day

FROM
(
SELECT user_pseudo_id, onboarding_study_group, count(event_name)/28 as sessions_per_day
FROM `oura-69259.analytics_164364211.v4_onboarding_study_session_start_data`


GROUP BY user_pseudo_id, onboarding_study_group
ORDER BY user_pseudo_id
)

GROUP BY onboarding_study_group
ORDER BY onboarding_study_group
```

Rate of ring wear

onboarding_study_ordering_ring_wear_data   LINK SHARING   HIDE EDITOR   FULL SCREEN

```

SELECT AVG(rate_worn) as rate_worn, onboarding_study_group
FROM
  (SELECT a.user_uid, a.rate_worn, b.onboarding_study_group

  FROM
    (SELECT user_uid, SUM(total_wear)/30240 as rate_worn

    FROM
      (SELECT o.user_uid, o.total_wear, o.datetime_local_rounded_1d as date, f.first_day,
TIMESTAMP_DIFF(o.datetime_local_rounded_1d, f.first_day, DAY) as day_number

      FROM `oura-69259.analytics_164364211.onboarding_study_ring_wear_data` as o, (SELECT user_uid,
MIN(datetime_local_rounded_1d) as first_day FROM `oura-69259.analytics_164364211.onboarding_study_ring_wear_data` GROUP BY
user_uid) as f

      WHERE o.user_uid = f.user_uid
      AND TIMESTAMP_DIFF(o.datetime_local_rounded_1d, f.first_day, DAY) <=20

      ORDER BY user_uid, date)


    GROUP BY user_uid
    ORDER BY rate_worn) as a, `oura-69259.analytics_164364211.onboarding_study_participant_list` as b

  WHERE a.user_uid = b.user_id
  AND rate_worn >= 0.1)
GROUP BY onboarding_study_group
ORDER BY rate_worn DESC
```

Appendix B

SQL queries for advanced analysis.

Sessions by device type

Generating sessions by device type — LINK SHARING — HIDE EDITOR — FULL SCREEN

```sql
SELECT user_pseudo_id, device.mobile_model_name, event_name, value.string_value as onboarding_study_group

FROM `oura-69259.analytics_164364211.master_onboarding_study_participant_data`

WHERE event_name = 'session_start'

ORDER BY user_pseudo_id

```

Sessions by country

Generating session counts by country — LINK SHARING — HIDE EDITOR — FULL SCREEN

```sql
SELECT user_pseudo_id, event_name, geo.sub_continent, geo.country, geo.city, value.string_value as onboarding_study_group

FROM `oura-69259.analytics_164364211.master_onboarding_study_participant_data`

WHERE event_name = 'session_start'

ORDER BY user_pseudo_id
```

Time in app by country

Generating engagement time by country — LINK SHARING — HIDE EDITOR — FULL SCREEN

```sql
SELECT user_pseudo_id, country, onboarding_study_group, SUM(engagement_time) as total_engagement_time
FROM
(
  SELECT user_pseudo_id, event_name, geo.sub_continent, geo.country, geo.city, value.string_value as onboarding_study_group, (SELECT value.int_value
  FROM UNNEST(event_params) WHERE key = "engagement_time_msec") AS engagement_time

  FROM `oura-69259.analytics_164364211.master_onboarding_study_participant_data`

  WHERE event_name = 'user_engagement'
  ORDER BY user_pseudo_id
)
WHERE user_pseudo_id IN (SELECT * FROM `oura-69259.analytics_164364211.onboarding_study_single_country_users`)

GROUP BY user_pseudo_id, onboarding_study_group, country
ORDER BY  onboarding_study_group, user_pseudo_id, total_engagement_time DESC

```