Hypnos — Developing a Sleep Tracking System through User-oriented Media Design

Gøran A. Slettemark

Supervisor Christoph Trattner

Thesis submitted for the degree of Master in Media and Interaction Design
Department of Information Science and Media Studies
University of Bergen

2019

Contents

1. Introduction

1.1. Motivation

Sleep tracking technologies has become a common selling point of fitness devices and smartwatches. Information that was previously solely used in the treatment of sleeping disorders is now in the hands of the consumer. These devices can automatically track the duration of sleep and infer the quality of sleep through activity tracking and biometric data.

Sleep is essential for good health, but many do no sleep enough or suffer from disrupted sleep [50]. Even slightly reduced sleep over several days may have ill effects, including higher instances of illness [2]. Getting enough sleep is also crucial for learning [6], and physical [41] and cognitive ability [24]. The Norwegian Institute of Public Health writes that one in three adults have weekly symptoms of insomnia. From 2000 to 2010, the percentage of adults with insomnia has gone from 11.9% to 15% [50].

Sleep tracking applications may state that they can help improve sleep. However, they often leave it to the user to understand the information the application presents, instead of providing actionable insights and learning opportunities.

Systems like *ShutI* offer a digital platform for the treatment of sleep disorders. Based on the established *cognitive behavioural therapy for insomnia* (CBTI), it allows the user to learn about healthy sleep habits via video and articles. It gives a personalised sleep schedule based on a sleep diary, which is collected via a website. While an evaluation calls this an "unguided treatment" [19], one could argue that this system becomes the guide. This sort of system has the opportunity to change the way many people interact with health professionals.

This thesis attempts to find a middle ground: a system that provides the ease of use that sleep tracking devices provide, but that can also provide advice on how to improve sleep habits. Such a system could allow for guided reflection, but also persuasion towards healthier sleep. Designing this system requires an understanding of how people can be persuaded to improve their sleep. It also requires knowledge about sleep itself in order to design recommendations that are effective and also safe.

1.2. Objectives

This project aims to better understand how users interact with sleep tracking systems in order to be able to propose design solutions. The thesis will refer to the following research questions:

- **RQ1**: How do users interact with existing sleep tracking systems?
- **RQ2**: What are the user motivations for using sleep tracking systems?
- **RQ3**: What functions do users find necessary in sleep tracking systems?
- **RQ4**: How can sleep tracking systems better support healthy sleep behaviour?

1.3. Contribution

This thesis will describe the use of user-oriented research methods that contribute a better perspective of what people want out of sleep tracking systems. These methods include interviews with users of existing sleep tracking applications, an online questionnaire about use and opinions of sleep tracking, and usability evaluations of sleep tracking interfaces. It will also describe the design of multiple prototypes, including low-fidelity prototypes and a functional high-fidelity prototype. The prototype itself is another contribution, as it can be used in further research and development.

1.4. Thesis outline

This thesis contains six chapters. Following this introduction, **Chapter 2** will review relevant literature regarding the biology of sleep, personal tracking, sleep tracking, and persuasive technology. **Chapter 3** will describe the methods used in this project, which includes *user studies*, *wireframe prototyping*, and the development of a *high-fidelity prototype*. **Chapter 4** will describe the results from the user studies. In Chapter 5 the final prototype will be presented. The final chapter – **Chapter 6** – will summarise the findings of this project from the perspective of the research questions outlined above. Following the main chapters is a collection of appendices, which contain supplemental documents related to the user research, and information on how to access the final prototype.

2. Background

This chapter will explain what sleep is from a biological perspective. It will then look at sleep deprivation and insomnia, and how sleep problems can be improved using *cognitive behavioural therapy*. Further, it will look into personal tracking using digital media before going into sleep tracking specifically. Finally, it will present examples of sleep tracking applications, as well as related prototypes and user studies.

2.1. Sleep

Sleep is an important restorative process for both mind and body. While the body is mostly inactive during sleep, the brain goes through several stages of activity, believed to be a way of consolidating memories [56].

How many hours of sleep a person needs varies, with age being one crucial factor. The U.S. National Sleep Foundation recommends seven to nine hours a night for adults, but six to ten "may be appropriate" [21]. For school-aged children (age 6 to 13), the recommendation is from nine to eleven hours. Numbers from the Norwegian Institute for Public Health indicate that Norwegian men sleep for slightly less than seven hours on average, while women sleep more than seven hours on average [50].

A lack of sleep has been shown to negatively influence short term memory and the ability to learn [6]. Purely physically, recovery from exercise is also affected [41]. Even slightly reduced sleep – less than 7 hours – over several days may have ill effects, including higher instances of illness [2].

The Norwegian Institute of Public Health writes that one in three adults have weekly symptoms of insomnia [50]. From 2000 to 2010, the percentage of adults with insomnia in Norway has gone from 11.9% to 15%, showing that it is an increasing problem. This report also notes that especially adolescents sleep less than recommended, getting only 6 hours and 25 minutes on weekdays.

2.1.1. Sleep stages

During one night of sleep, a person will cycle between the two main stages of sleep: rapid eye movement (REM) sleep, and non-rapid eye movement sleep (NREM). A complete sleep cycle takes approximately 90 minutes, with the proportions of NREM to REM sleep shifting throughout the night. The first cycles have more NREM sleep, while later cycles will have more REM sleep. Both stages of sleep are specialised: deep NREM sleep is more restorative in terms of the body, it is where the body secretes the most growth hormones [18], while REM sleep is theorised to be important in the storage of memories [42].

2.1.2. The need for sleep

Two physiological factors influence the need for sleep, and these are called *homeostatic* and *circadian* factors [18].

Homeostatic factors are related to how long it has been since sleeping last. One such factor is the level of the chemical adenosine in the body. When awake, levels of adenosine in the body rises, which leads to sleepiness. During sleep, the level of this chemical decrease and the need for sleep also decreases. Sleep deprivation can lead to a build-up of this chemical, which will lead to daytime drowsiness, as the body will want to sleep in order to reduce the levels of this chemical. Many are aware of the impact of caffeine on sleepiness. Caffeine works by blocking adenosine receptors in the brain, meaning that it will make someone feel less tired. However, adenosine will continue building up in the body. With repeated use of caffeine, the brain will build more adenosine receptors, making the person require more caffeine for the same effect.

The **Circadian factor**, or the biological clock, is determined by external factors, with the most important being light exposure. Light help our brain set the biological clock by activating a part of the brain called the suprachiasmatic nucleus. This part of the brain is responsible for keeping track of time and sends signals to other parts of the brain that regulates hormone production and other biological functions. [18]

Circadian rhythms are the cause of jet lag, which is when the brain's understanding of time and the time in the surrounding location is mismatched [58]. Other lights, like computer screens, are possibly disrupting the circadian system in the same way, as the light in most computer and phone screens are the right wavelength to stimulate the receptors in the brain the same way sunlight does [9].

Chronotypes

The natural circadian rhythm of a person is also called their *chronotype*. The chronotype changes with age, with young children usually being *early* chronotypes, and teenagers *late* chronotypes. An early chronotype will naturally wake up earlier in the morning, while a late chronotype may slumber past noon. In the mid-twenties, the chronotype usually stabilises until old age [45].

Chronotypes can be measured by mid-sleep time, which is the midpoint between when a person goes to sleep and when they wake up. Issues like social jet lag, the mismatch

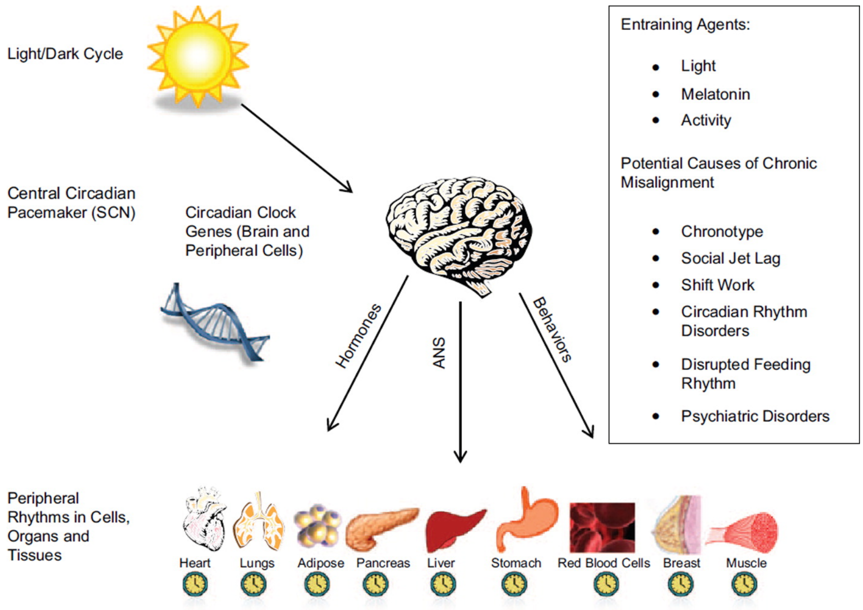


Figure 2.1.: Representation of central and peripheral circadian rhythms. Reprinted from Baron and Reid [3]

between social hours and chronotype, can be identified by looking at the difference between the midpoint of sleep on working days and days off [59]. Late chronotypes are more likely to be affected by social jetlag, as work start times will restrict their time spent sleeping. Wittmann et al. [59] suggest that they may consume more caffeinated drinks, nicotine and alcohol as a coping mechanism. Late chronotypes were also more likely to report a depressed mood.

Social jetlag may result in a *circadian misalignment*, which is the misalignment of one of many functions of the body concerning the biological clock. This type of misalignment may manifest as sleep disturbances or daytime sleepiness. The consequences can include changes in dietary behaviour and appetite regulation, glucose regulation, and mood, which may again increase the risk for cardiovascular disease, diabetes and psychiatric conditions. [3]

2.1.3. Measuring sleep quality

Sleep quality can be measured via activity in the brain or the body, using techniques such as EEG or actigraphy. These measurements are used to indicate how much time is spent in different sleep stages. Not spending enough time in all sleep stages can be considered disordered sleep. Sleep quality can also be measured subjectively, often using a scale from one to five. A 1994 study showed that subjective sleep quality is closely related to objective measures of sleep, especially *sleep efficiency* [1].

One example of subjective sleep quality measurement is the *Pittsburgh sleep quality index* (PSQI) [8]. It is meant to assess the sleep quality of the previous month, and it consists of 19 items, grouped into seven components, which are:

- Sleep quality
- Sleep latency
- Sleep length
- Habitual sleep efficiency
- Sleep disturbance
- Use of sleeping medication
- Daytime dysfunction

These components are given a score from between 0 and 3, and together they form a total score. The higher the score, the worse the sleep quality.

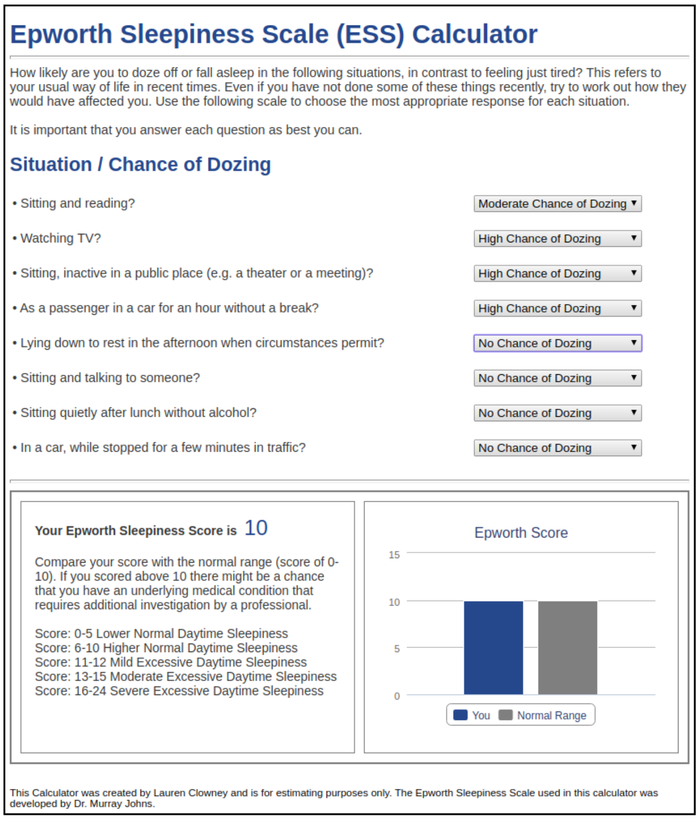


Figure 2.2.: Screenshot from an online Epworth Sleepiness Scale calculator. From http://www.calcz.com/apnea/calc.html

Another scale is the Epworth Sleepiness Scale (ESS) [25]. It is an assessment of daytime sleepiness and asks the user to rate to which extent they are likely to "doze" in eight different situations using a scale from one to four. A sum of below or equal to ten is considered normal [25], while a sum above ten can be an indication of a sleep disorder.

2.1.4. Sleep deprivation

Sleep deprivation, or insufficient sleep, can have several causes. It can be caused by insomnia, where a person may have problems with falling asleep or waking up too early. Another cause may be that a person does not give themselves enough time to sleep, because they prioritise other activities, or do not believe they need more sleep. Stressful events can cause a vulnerability to insomnia [30]. External factors like temperature, light, and noise level can also play a part, by reducing the quality of sleep [36].

Consequences of sleep deprivation

Experiments have shown that regularly sleeping less than seven hours a night can lead to "significant daytime cognitive dysfunction", including reduced attention and working memory [2].

Chronic sleep deprivation leads to an upregulation of appetite and lowers energy expenditure. Together with altered glucose metabolism, these factors can likely increase the risk for type 2 diabetes [26]. Sleep deprivation is also related to a reduction in the immune response [23]. Since 1924, experiments have shown that sleep deprivation has a negative effect on memory and learning [24].

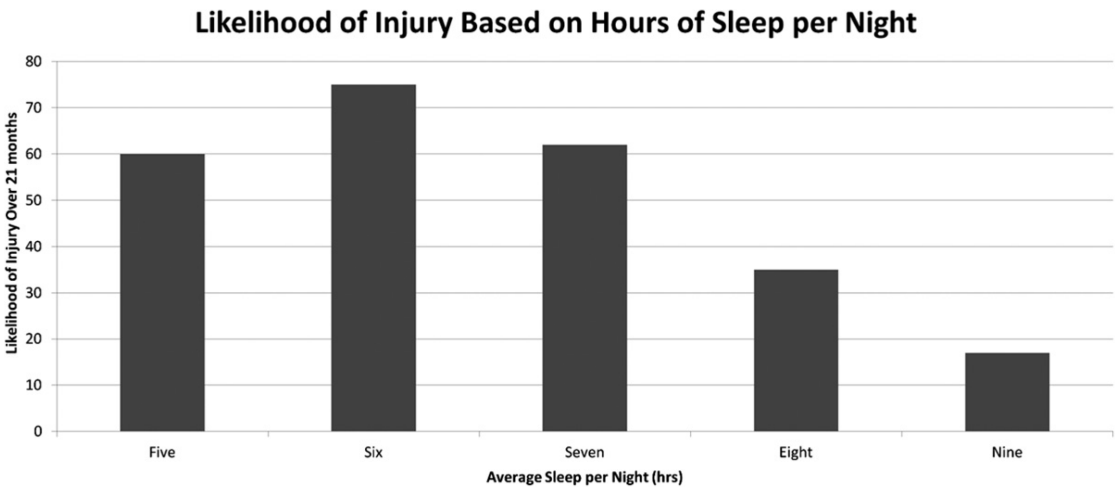


Figure 2.3.: Likelihood of Injury Based on Hours of Sleep per Night. Reprinted from Milewski et al. [29]

Researchers have also shown that sports injuries are more likely when the athlete has been deprived of sleep. Those who slept for less than 8 hours a night on average was 1.7 times more likely to have needed treatment for an injury, compared with those who slept for more than 8 hours [29]. Another study looking at recovery from exercise showed reduced power output in athletes subjected to one night of sleep deprivation [41].

2.1.5. Insomnia

Insomnia can manifest in several ways. It can be difficulties with getting to sleep, maintaining sleep or waking up early without being able to return to sleep. Insomnia is usually seen together with *sleep opportunity*, the time a person sets aside for sleep.

Insomnia can be **transient**, lasting for less than a week; **acute**, which is consistent sleep problems for less than a month; and **chronic**, when problems have lasted for longer than a month. It can be *primary*, which means that there is no apparent disease causing the inadequate sleep, or *secondary*, meaning it is caused by disorders like depression, heart disease, and similar systemic conditions [48].

Treatment of insomnia

The preferred therapy for chronic insomnia, after other possible causes are ruled out, is *cognitive behavioural therapy* (CBT) [44]. It has been shown to be as effective – if not more effective in the long term – as common sleep drugs [51]. Cognitive behavioural therapy for insomnia (CBT-I) includes practising *sleep hygiene* and *stimulus control*, and sometimes *sleep restriction* [31].

Sleep hygiene is keeping the sleep environment free from distractions, such as TVs, computers or similar. The bedroom should be dark and quiet. Keeping a consistent bedtime and wake-up time is also part of good sleep hygiene. The person should only go to bed when they are tired and should avoid staying in bed if they are unable to get to sleep. [31]

Stimulus control is about limiting stimulating activities before bedtime. This can mean avoiding late meals, alcohol or caffeine, certain medications, stressful situations or strenuous exercise. It may also include beneficial activities such as exercising during the day or relaxing activities before bedtime. [31]

Sleep restriction is often a part of CBT-I. Limiting the amount of time a person spends in bed to the time they spend sleeping can help restore the ability to fall asleep or stay asleep. The sleep time is then adjusted based on *sleep efficiency*, the percentage of time the person spends sleeping while in bed. If it is above 90%, it can be increased by 20-30 minutes, and if it is below 80%, it may be decreased by the same amount. [31]

Keeping a sleep journal is a common practice for patients undergoing CBT-I. It is used in order to calculate sleep efficiency, as well as allowing the patient to reflect on their sleep. In an example from *Nasjonal kompetansetjeneste for søvnsykdommer* [5], the patient is asked to fill out ten fields every day:

1. Subjective rating of daily functioning (1 - 5)
2. Time and duration of naps.
3. Sleep medication or alcohol intake
4. Bedtime and the time of attempted sleep
5. How long it took to fall asleep
6. Number of awakenings
7. Total time spent awake during the awakenings
8. Wake-up time
9. Time getting out of bed
10. Subjective rating of sleep quality (1 - 5)

The next section will look at the modern cousin of journal keeping: personal tracking using digital media.

2.2. Personal tracking

Personal tracking, self-tracking, or the quantified self [55], is the concept of keeping track of data related to oneself. This data can be location data, usage data, activity sensors, diet tracking, and even blood sampling.

Some tracking is *manual*, like most diet tracking applications, where the user has to enter data themselves. Various sensors can make *automatic* tracking possible. One such example is wearable activity trackers that can measure activity through an always-on accelerometer. Some devices, like a GPS tracker, can track data automatically after being manually activated.

Many commercial products that are health or wellness related are using tracked data for motivation or for guiding users towards a goal. *Gamification*, which is the use of elements from video games, can be used to allow the user to earn points or badges for completing tasks, possibly persuading the user to do things they would not otherwise.

2.2.1. Categories of personal tracking

Rooksby et al. [47] identified five styles of tracking: directive tracking, documentary tracking, diagnostic tracking, collecting rewards, and fetishised tracking.

Directive tracking could also be described as goal-oriented tracking. This style of personal tracking is oriented around a goal, such as weight-loss or training for a sports event. Here, the tracking system is acting as a tool to support the user reaching their goal. The authors noted that "[s]ometimes the goal would come directly via the tracker itself; most pedometers for example suggest appropriate goals."

Documentary tracking is tracking more for the sake of documenting activity. The reasons for this type of tracking were many, but the authors mention the user's desire to compare themselves to other people as a primary motivator. Other users of this type were motivated by wanting to get credit for their activity level. Some users used this style of tracking intermittently: "once some people had a sense of how many steps it was to work, or what speeds they were doing on a regular cycle ride, then it was no longer necessary to record them." In general, the authors note that documentary tracking is not a perpetual affair; users who lean towards this style of tracking do not collect data simply to accrue them. The authors mention sleep tracking as an exception to this: "Sleep tracking was perhaps the only area where logs were being accrued in isolation from specific goals." Documentary tracking can also transition into goal-driven tracking, if the user discovers a goal during their documentary effort.

Diagnostic tracking is described as tracking where the user is "[...] looking for a link between one thing and another." One participant had tracked his diet and medications to diagnose a stomach problem. Another had acquired a sleep tracker to find out why he was tired in the mornings. The author noted that: "Diagnostic tracking was done over a period of time but did not need to carry on indefinitely, just until the person was satisfied that they have an answer."

Some used tracking in order to collect rewards. These rewards could be either gamified elements within the application or actual monetary rewards. Examples of this kind of tracking could be getting medals on Strava, or badges for reaching a particular step goal with a FitBit. The now-defunct application Pact paid users who reached their set goals by charging those who did not.

Finally, the authors briefly describe **fetishised tracking**. By this, they mean tracking motivated by a "purer interest in gadgets and technology." In other words, tracking where the tracking technology is the primary motivating factor.

2.2.2. Long term tracking

Fritz et al. [17] interviewed 30 users of fitness devices (Nike FuelBand, and an early non-wristband Fitbit) that had used the device for at least three months.

They found that users had become attached to their devices, but some users were less excited about their device after some use. Users had become more aware of how their behaviour influenced the score or step count the device provided. Users reported both *immediate impacts* and *durable changes* that had occurred because of using the device. Users were motivated by reaching their self-subjected goals, but also by *system rewards* that is rewarded by completing tasks. When it came to social features, some users had reservations due to privacy, and others were just not interested in sharing their data. Still, those who used the social features found it motivating.

The authors suggest that designers should consider if their technology is meant to function as a *gateway technology* or if it is meant to support behaviour in the longer term. They suggest that systems could offer better motivation for maintaining behaviour, as it is important for behaviour change. For some of their participants, step counting had become less relevant as they found other fitness activities. They suggest that systems should be able to be "augmented or evolved over time", by supporting new activities. Rewards should also be able to change over time, and the authors would like to see a greater variety of rewards, perhaps including real-world rewards.

2.2.3. Social sharing

In a 2012 study, Munson and Consolvo [34] explored how various motivations, including goal-setting, rewards, and social sharing, affected users. The study used two related applications: Goal-post and Goal-line, which are both activity tracking systems. In both systems, the user could enter how much time they spent with different activities such as cardio, strength, and walking. The user could also set both a primary goal and secondary goals. The difference between the two systems is that Goal-post included a social sharing feature, that allowed the user to post either an activity or a completed goal, to their Facebook newsfeed.

The study suggested that goal-setting can benefit from including both secondary and primary goals. The authors write that "simultaneous goals helped participants stretch beyond what they would otherwise have done and gave them a fallback during busy weeks that helped them remember to do at least some exercise." Their participants were also positive towards reminders, and they suggest that the reminder should include progress towards a goal. The participants did not find digital rewards motivating. The authors suggested this was due to the implementation of the rewards not being tied to progress in a meaningful way.

Participants had problems with social sharing: some did not want to share too much or too often, others did not find support when they shared. The authors suggest that making the information that is shared more attractive could be a solution to this. They also suggest that sharing could be done to a network of invited peers instead of to the more public space of Facebook.

2.3. Sleep tracking

Sleep tracking can be simply described as the act of keeping track of sleep. For medical treatments, sleep diaries have been used to measure the length of sleep, subjective sleep quality and potential disruptors of sleep. Traditionally this was done using pen and paper, but more recently sleep tracking has become a selling point of wearable devices and smartphone applications. Wearable sleep tracking systems, as part of medical treatments of sleep disorders, has also shown promise [49].

Wearable sleep trackers use sensors to measure sleep so that the user does not have to remember when they went to bed, and when wake up. Some also attempt to measure sleep quality, showing the user how much time they spend in different sleep stages.

Smartphone sleep tracking system does also exist, often requiring the user to place their phone in or near their bed. These systems usually rely on the accelerometer or microphone of a smartphone. A regular feature of a smartphone system is the "smart alarm", that attempts to wake the user up when their sleep is lighter. There are also applications that more closely mimics the traditional sleep diary, as will be described in the next section.

2.3.1. Examples of sleep tracking products

This section will describe a selection of existing commercial sleep tracking systems. These systems were selected as representatives of different types of sleep tracking systems. The systems described include *SHUTi*, a treatment system for insomnia, *Fitbit*, a wearable fitness band, *Oura*, a wearable ring that focuses on heart rate, *Sleep as an Droid*, a smartphone application, and *SleepTown*, another smartphone application that focuses on social sharing.

SHUTi (Sleep Healthy Using the Internet) is a digital treatment system based on *cognitive behavioural therapy for insomnia* (CBT-I). Central to this therapy is learning about what factors can impact sleep, which is usually done with a therapist. Sleep diaries are commonly used, both as a tool for reflection, but also to calculate *sleep efficiency.* Sleep efficiency is the time a person spends asleep divided by the time they spend in bed. Based on this efficiency, a *sleep window* is prescribed. The sleep window is usually the time the persons spends sleeping.

SHUTi contains a digital sleep diary where the user can enter their bedtime, the time it took them to fall asleep, number of awakenings, when they woke up and got out of bed and a rating of their sleep from *"very deep"* to *"very light"*. The diary also asks about the duration of any naps during the day, and whether they drank alcohol or took any sleep aids. Finally, the user can enter a personal note. Except for the personal note, this is all done through drop-down menus.

SHUTi has several learning modules that the user unlocks as they keep using the system. These modules are made up of video and text that is supposed to help the user learn about each subject. After completing each module, the user is asked questions related to their progress and what they found challenging, which is another effort to push the user towards self-reflection.

When compared to purely educational websites, users of SHUTi had a significantly better outcome on measurements of insomnia. This outcome was also present in a six-month follow-up. [19]

Fitbit is one of the most popular fitness wristbands. Most of Fitbit's models have sleep tracking built-in, but the technology that is used to measure sleep varies. Fitbit focuses on automatic tracking, which means the user simply has to wear their Fitbit, and it does the job of figuring out when they go to bed and wake up. The accelerometer is used to detect when the user is still, which can be used to detect sleep and sleep quality. Newer devices also use heart rate, and some have a blood oxygen sensor that may be used in the future to enhance the sleep quality estimate further.

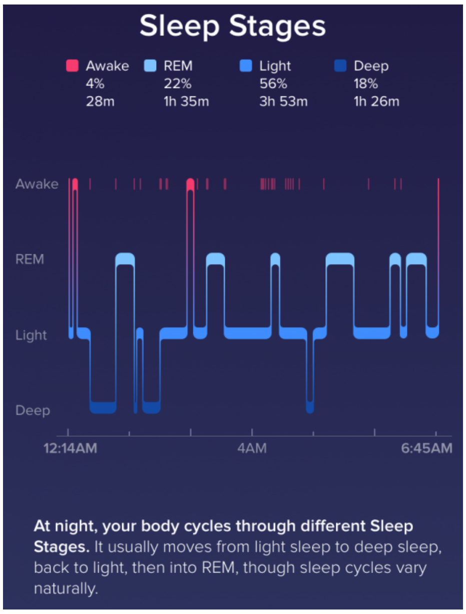
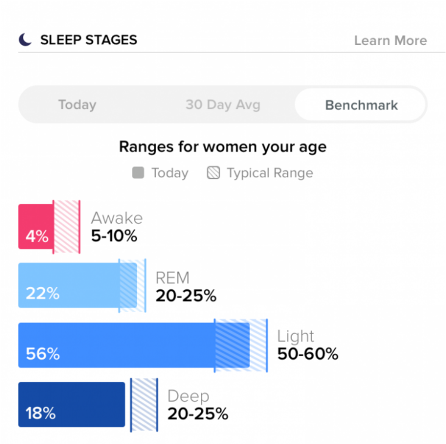


Figure 2.4.: How Fitbit represents sleep stages. From the Fitbit blog [15]

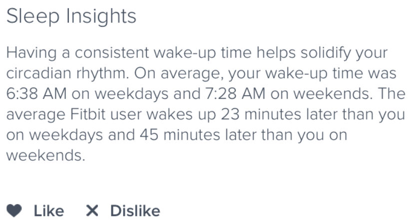
Fitbit's sleep stage tracking uses motion detectors and heart rate data to detect time spent in different sleep stages. According to their categories, the user can be either *Awake*, or in *REM*, *Light Sleep*, or *Deep Sleep*. The time spent in each stage is shown to the user in several ways. It is displayed as a chronological graph (see figure 2.4) that shows which stage the user was in at that time. It is also displayed as a bar graph with percentages, showing how much time was spent in each stage. Worth noting is also the *Benchmark* feature, which shows how the user's time spent in each stage compares to a *typical range* of users of the same gender and age.

With the *Sleep Schedule* feature, Fitbit allows the user to define a sleep goal by the number of hours, as well as bedtime or wake-up time. If either bedtime or a wake-up time target is not set, the Fitbit app will estimate based on recent trends. Based on this, the user can get bedtime reminders and earn stars when they reach their goals.

With *Sleep insights* Fitbit attempts to give the user personalised information in the form of text snippets. According to their website, it can "[...]help you learn more about your sleep stats and how your trends compare to those of the same age and gender." [52].



(a) Fitbit's sleep benchmark



(b) Fitbit's sleep insights

Figure 2.5.: (c): From the Fitbit blog [15], (b): from a Fitbit help article [14]

Figure 2.6.: Fitbit's sleep score. Screenshot by Jason Cipriani/ZDNet [12]

Fitbit has recently experimented with a *sleep score* feature beta [12]. This feature gives each night a score from 0 to 100. This total score is made up of another three scores: one for sleep duration, one for *sleep depth* and one for *revitalization*. While sleep duration and sleep depth are just giving a score to data that was already there, the revitalization score is something new. From the description: "This category explores how restorative your sleep is by detecting breathing disturbances, and comparing your sleeping heart rate to your daytime stats to measure how refreshed you should feel in the morning".

By focusing on more physical data, Fitbit could be better able to help those who have problems with sleep apnea or other physical sleep disturbances.

Oura is a wearable ring that connects to a smartphone. While it has a lot in common with Fitbit, the Oura ring is interesting because it has a different focus in displaying data to the user. One of the selling points is its heart rate sensor, which can track *heart rate variability*, a measurement of the autonomic nervous system that can respond to exercise, food, sleep and stress.

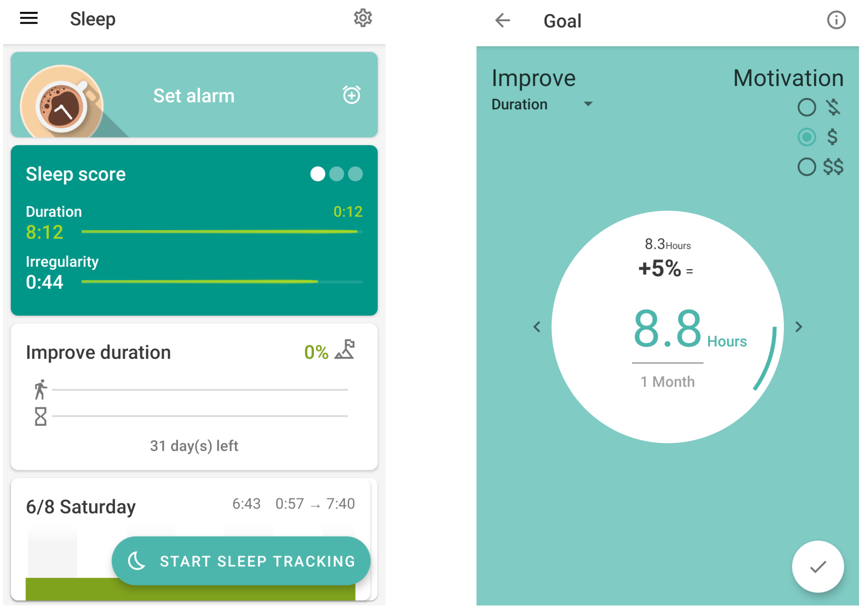
(a) (b)

Figure 2.7.: Oura's sleep interface. From the Google Play Store [39]

Like the Fitbit, the Oura automatically tracks sleep duration, and also sleep stages. It displays the sleep stages in much the same way as Fitbit (except it has no benchmark). However, the focus on heart rate shows up again here: in addition to the sleep stages graph, it displays the heart rate during sleep.

Oura uses a score between 0 and 100 to grade the user's sleep with seven *sleep contributors* [22]. These include the amount of time slept, sleep efficiency, *restfulness*(may be renamed disturbances), REM sleep, Deep sleep, *Latency* and timing. The focus on timing as a part of the score is an exciting difference to Fitbit. The Oura rewards consistent sleep timing with a higher score, whereas Fitbit seems to focus on measurable data.

Sleep is also reflected in the overall readiness score that the Oura app gives. Along with other measurements like the amount of exercise, resting heart rate, body temperature and *recovery index*, two sleep-related contributors are found. The first is simply the previous night's sleep score. The other is called *sleep balance* and is supposedly based on a comparison of the past two weeks of sleep with a longer-term baseline and recommendations for the user's age [22].



(a) (b)

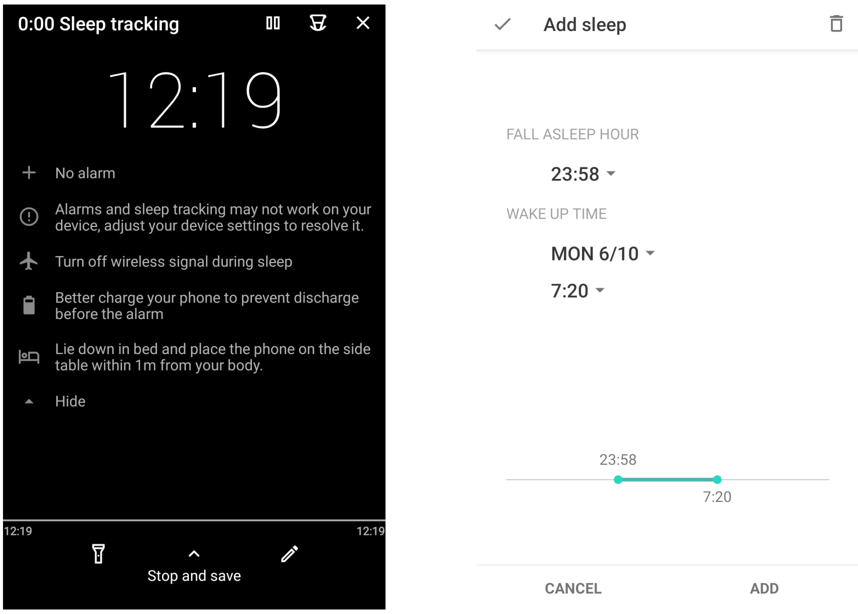
Figure 2.8.: Screenshots from Sleep As an Droid's home screen (a) and goal-setting component (b)

Sleep as an Droid is a sleep tracking application for Android smartphones. It can track sleep in several ways. One way is using the accelerometer, where the user places their phone in their bed (while connected to a charger). An alternative to this is *sonar* tracking, which relies on ultrasound emitted by the speaker and retrieved by the microphone. The upshot of this is that the phone can be placed next to the bed instead of in the bed. Both of these types of tracking have to be manually started and stopped by the user. Sleep as an Droid also has support for several wearable devices as an alternative data source. It also has manual sleep tracking, and automatic tracking based on phone use.

Being a smartphone app allows Sleep as an Droid several opportunities that most wearable devices cannot currently match. The application boasts the following features:

- record sleep talk and snoring
- attempt to counteract snoring by playing certain sounds
- try to induce lucid dreams by playing certain sounds
- play music to help the user fall asleep
- attempt to wake the user up when they are in a specific sleep stage

A smart alarm is a selling point of several similar smartphone applications, where the idea is that waking up in the right sleep stage may make it more comfortable or more



(a) (b)

Figure 2.9.: Screenshots from Sleep As an Droid while tracking (a) and the alternative manual input (b)

natural. Wearable devices cannot currently do this because they do not process the sleep data as it is tracking; the data processing takes place after syncing with a phone.

In terms of data, Sleep as an Droid is also wide-reaching. The user can view much of the same information as on a Fitbit, and it has a sleep score system. It also makes simple correlations, either between time spent asleep and estimated sleep quality, or bedtime/waketime and subjective sleep quality.

SleepTown is an example of a gamified smartphone sleep tracker. With a focus on reaching user-defined goals, such as sleeping eight hours and getting in bed before 11, it rewards the user with constructing virtual buildings in their very own *sleep town*. As they complete goals, their town grows, and with consistency, they get the opportunity to build buildings with a higher rarity. The user can move buildings around in their town and spend coins on cosmetic items. Users can also compare their towns with friends, and cooperate to build special buildings.

SleepTown does not track sleep stages or activity during sleep; it focuses on rewarding not using the phone during the night. When the user commits to going to sleep, the application locks down the phone. The user then has to *"collapse"* their building before being able to use their phone during the night. It still keeps track of the amount of sleep, which can be viewed as graphs similar to Fitbit or Sleep as an Droid.

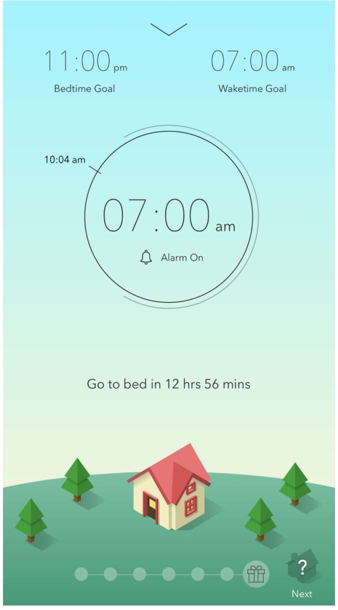


Figure 2.10.: SleepTown's main interface. From the SleepTown website [54]

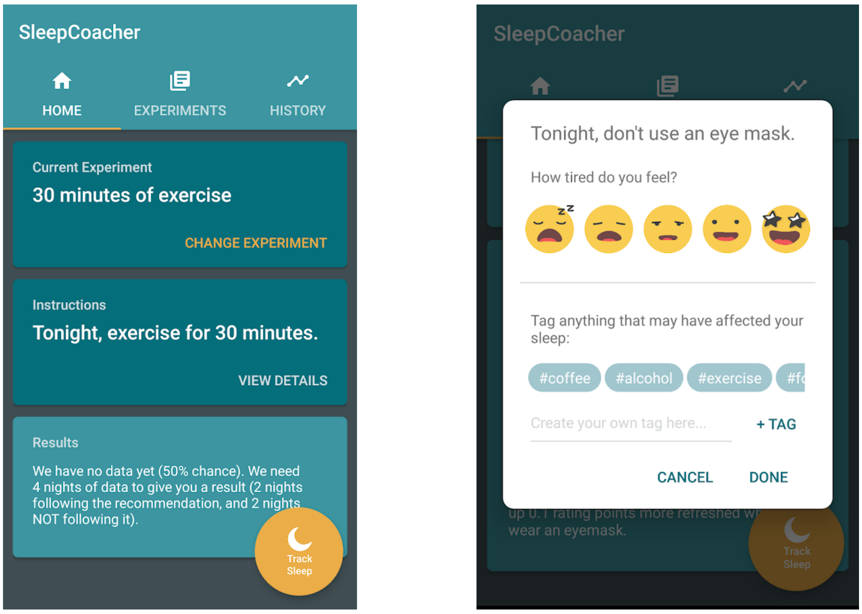
2.3.2. Sleep tracking prototypes

This section will look at sleep tracking prototypes coming out of academic institutions. Most are research prototypes, developed as part of projects in either human-computer interaction or interaction design.

ShutEye [4] With ShutEye, Bauer et al. created a dynamic smartphone wallpaper that attempts to promote healthier sleep habits. The wallpaper displays different activities that can impact sleep, like eating, napping, exercise, alcohol and relaxation. The activities are represented as a series of bars. The thickness of each bar indicates if the activity is allowed or discouraged. In addition to the wallpaper, there is a companion app where users can customise the times associated with each activity, and get a description of how each activity can impact sleep.

A four-week field test indicated that the application could make users more aware of how these activities impacts sleep habits. It showed that users who did not see the recommendation as rules but more as suggestions had more success. Some users understood an activity that was marked "allowed" as "encouraged", which was not always the case. The authors write that this could be remedied with having three levels (discouraged, allowable, encouraged) instead of just allowed and discouraged.

As users responded well to being able to read more about the recommendations, the authors suggest that similar applications should provide both a summary and a link to an authoritative source. Users also were interested in suggestions to how they could meet the recommendations. The authors write that future systems could present information like alternatives to caffeinated drinks, relaxation techniques, and also recommend exercise and meals that is appropriate for the time of day.



(a) (b)

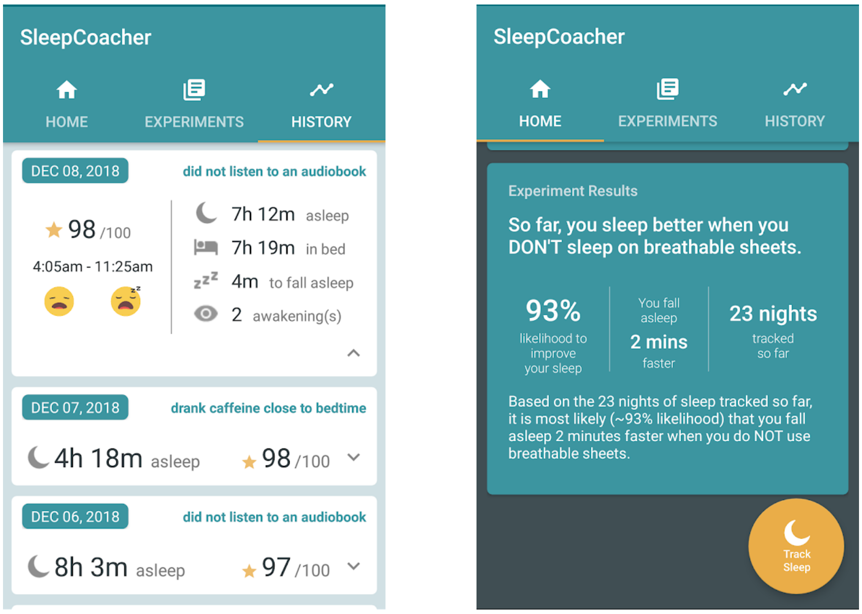
Figure 2.11.: Sleepcoacher on Android. From the Google Play Store [53]

Sleepcoacher [13] is a smartphone-based sleep tracking application that guides the user through self-experiments. The user can select between several sleep-related recommendations such as using ear-plugs, wearing socks at night or meditating before sleep. Then the application guides the user through an ABAB-experiment, where A is not following the recommendation, and B is following the recommendation. Each phase has several data points, which in this case is 3-5 nights of sleep.

The author describes SleepCoacher as a closed loop that tries to "[...] determine whether a behavior change occurred and yielded improvements in targeted aspects of sleep [...]". The application relies on sensors such as an accelerometer and microphone. It also asks the user for self-reported sleep ratings. As most smartphone sleep tracking applications, the tracking has to be manually started before going to sleep.

A user study showed that those who adhered to the sleep experiments were likely to improve their sleep: "There is improvement in 13 of the 16 cases when adherence rate is higher than 60%, but only 9 of the 18 cases with [a] rate lower than 60% improved. Target sleep variables were improved in all 7 of the cases when adherence was higher than 80%."

Users responded best to recommendations that were easier to follow: "When users found



(a) (b)

Figure 2.12.: Sleepcoacher on Android. From the Google Play Store [53]

the effort- or time-cost of following a recommendation to be low, many were happy to follow recommendations". The analysis found correlations between time slept and rating, but not between rating and bedtime, which the author claim is evidence of the need for personalisation in this type of system.

SleepTight [11] With SleepTight, Choe et al. compared a sleep tracking system using a widget on a smartphone home screen, to a sleep tracking system without a widget. Central to the design of SleepTight is the ability to track *contributing factors* such as meals, exercise, and caffeine intake. The user could either do this using the widget or through the application itself.

In a four-week deployment study with 22 participants, they found that the version with the widget had higher compliance rates (92%) than the one without (73%). Users also reported that "[...] the widgets improved information access and encouraged self-reflection".

Data collected during the study showed that most users did not track contributing factors in real-time, but more often, they did so in retrospect. Users often preferred to do this in the evening, towards bedtime.

The authors speculate that SleepTight could be further developed to be able to give actionable advice. The focus should be on identifying routines and anomalies: "once a self-monitoring tool identifies people's routines, it should distinguish routines from anomalies and encourage people to collect anomalies. Rare events are valuable data points."

2.3.3. Sleep tracking user studies

This section presents studies of users of sleep tracking systems, as well as a study looking at how potential users relate to sleep.

Users of sleep sensing systems

Ravichandran et al. [43] conducted a series of investigations into sleep tracking, including a survey, interviews with users and medical experts, as well as a qualitative analysis of both application and device reviews. These investigations focused on *sleep sensing* technology and excluded self-reported tools such as sleep diaries.

They found that users of sleep tracking may develop "broken mental models" of how a sleep tracker works and what it can sense. Users were also frustrated with the lack of transparency in sleep sensing systems. Users found it distracting that the feedback from the sleep sensing systems focused on "unconscious aspects of sleep, such as time in sleep stages", instead of aspects they could more easily influence. The authors suggest that sleep tracking systems should focus on "duration, timing, and making connections to modifiable behaviours and sleep hygiene".

Based on their findings, they propose a series of design recommendations, that says a sleep sensing system should:

- Include subjective sleep quality assessment
- Contextualize sleep quality with journaling
- Focus on actionable feedback
- Give feedback in ranges, not single-point values
- Increase transparency in formulae and algorithms

Potential users

Another study of sleep tracking technology was done by Choe et al. [10]. They conducted interviews with sleep experts and surveyed and interviewed potential users.

Their survey focused on sleep behaviour and suggested that respondents had inconsistent sleep schedules, and slept longer on the weekends than on weekdays. 27% reported not having a regular sleep habit and relied on frequent naps. Many respondents were aware of the importance of a healthy sleep schedule and had also attempted to maintain a sleep routine. External factors like jobs and school often determined sleep schedules.

Worries, fears and work or school-related stress were commonly reported as sleep disruptors, more common than environmental factors like temperature and loud noises. Other disruptors were late-night caffeine, alcohol, playing video games, and surfing the web.

Increasing or decreasing temperature was the most popular sleep aid, followed by physical activity. People also reported using music, meditation, mental exercises, warm milk, tea, or alcohol, warm baths, rain sounds, and reading. The authors write that some had misconceptions about what they should do to sleep better, and that "there is room for technology to help bridge the gap by providing accurate and customized sleep hygiene information".

62.6% answered yes or maybe when questioned if they would be interested in using technology to help them sleep. These participants were significantly more likely to have had sleep-related problems in the past.

When asked about their interest in features, participants wanted sleep recording to be automatic, an assessment of sleep quality, recommendations for optimal sleep conditions and long-term sleep trends. Daylight simulation and help with maintaining regular routines were also mentioned. The least popular features included those that required manual input and sharing data on social networks.

Barriers to sleep tracking

Liang and Ploderer [28] interviewed 12 users who had been tracking their sleep using a Fitbit device. They found that sleep tracking raised awareness of sleep patterns, but the participants did not report having improved their sleep. Only three participants had changed their habits in order to improve sleep. Six participants said that the sleep tracker did not help them improve their sleep.

The authors identified the following barriers:

Not knowing what normal sleep is: Users were not sure if their sleep was healthy, and the application did not provide reference points that the user could compare themselves with. The authors suggest that the application could compare the user to individuals of the same age, or establish a reference for the user in another way.

Users also had a problem with the accuracy of the sleep tracking system. One issue was with it registering sleep when they were watching TV. These issues, in turn, made it difficult to compare one night of sleep to another, as the user was unsure of the correctness of each night.

Not identifying reasons for sleep problems: Some users wanted more help in connecting the dots between their sleep and contributing factors like exercise and lifestyle. The Fitbit application did track both steps and diet but did not make any connections between them. Some users had stopped tracking data that required manual input, such as diet, because they felt that it required too much effort. Others wanted to track other factors, but the application did not support it.

Not knowing how to act: The weekly reports that Fitbit provided did not help users to decide what actions to take in order to improve their sleep. The goal-setting functionality in the Fitbit application did not help users when it came to sleep since they could not easily increase their time spent sleeping. To remedy this, the authors suggest offering personalised recommendations, that takes the user's schedule, age and gender into account.

2.4. Persuasive technology

This chapter begins by describing the concept of *persuasive technology*. It will further look at how a digital medium can function as a social actor, and at the different strategies, such an actor can employ.

2.4.1. The persuasive computer

In a 1998 paper, Fogg [16] defines a persuasive computer as "an interactive technology that changes a person's attitudes or behaviors". For it to be true persuasion, it has to have the intent of changing behaviour. He writes that since computers do not have an intent itself, it is the creator, the distributor, or the adopter of the technology that makes the technology persuasive.

He further describes the **functional triad**, three ways of understanding an interactive technology: as a *tool*, as a *medium*, and as a *social actor*.

A **tool** is something that provides the user with abilities or that make a task easier to accomplish. Fogg writes that tools can be persuasive by reducing barriers, making a new behaviour seem achievable, allowing for informed decisions, and by shaping a person's mental model.

As a **medium**, the technology conveys information, such as text, graphs, video or 3D graphics. A medium provides experiences, learning, and insights that can be novel or thought-provoking. Certain experiences can help the user understand cause and effect, while others may motivate the user.

When technologies attempt to emulate a human or something a human would relate to socially, it can be considered a **social actor**. A social actor attempts to create a relationship with the user and uses existing social dynamics to do so.

2.4.2. Persuasive strategies

As part of a study where they wanted to investigate the relationship between personality and perception of persuasive technologies, Halko and Kientz [20] came up with the following set of persuasive strategies:

> 1. Instruction Style
>   - Authoritative: Uses an authoritative agent, such as a drill sergeant or strict personal trainer, to instruct the user on how to meet their fitness goals.
>   - Non-Authoritative: Uses a neutral agent, such as a friend or peer, to encourage the user to meet their goals.
> 2. Social Feedback
>   - Cooperative: Uses the notion of users cooperating as a team with friends or peers to complete their fitness goals.
>   - Competitive: Uses a strategy of competing against friends or peers to "win" a competition.
> 3. Motivation Type
>   - Extrinsic: Uses external motivators, such as winning trophies, as a reward for conducting healthy behaviors.
>   - Intrinsic: Uses internal motivators, such as feeling good about one's self or feeling healthy, to motivate healthy behaviors.
> 4. Reinforcement Type
>   - Negative Reinforcement: Removes an aversive stimulus (e.g., turns a brown and dying nature scene green and healthy) as the user conducts more healthy behaviors.
>   - Positive Reinforcement: Adds a positive stimulus (e.g., adds flowers, butterflies, and other nice-looking elements to any empty nature scene) as the user conducts more healthy behaviors.
>
> – Halko and Kientz [20]

When comparing usability tests of storyboards that were meant to invoke different persuasive strategies with a personality test, they found several significant correlations [20]. For example, a person that scored high in *openness* rated their likelihood of using a system that used an *authoritative* persuasive strategy higher. These findings suggest that tailoring persuasive strategies to personality may be a fruitful undertaking.

2.4.3. Reflective technology

Munson [35] has written about mindfulness, reflection and persuasion in personal tracking. He compares *persuasive technology* with his idea of *reflective technology*. He defines persuasive technology as systems that "[...] push people who interact with them to behave in certain ways, without those people choosing behavior change as an explicit goal".

Unlike Fogg, he writes that persuasive technologies may or may not be intended to be persuasive and that most systems that allow for user choice have some form of persuasion built-in.

Reflective technologies can also be persuasive, but they intend to enable the user to reflect on their behaviour. A reflective system should attempt to surface data neutrally in order to allow the user to conclude themselves. Munson suggests that *context sensing* could be a part of this; in different contexts, a system could focus on more or less on reflection or persuasion, depending on the user or defined goals. As shown by Halko and Kientz [20], personality could be one such factor.

2.4.4. Aesthetics in persuasive health applications

A study by Oyibo, Adaji, and Vassileva [40], looked into the relationship between *perceived credibility* and visual aesthetics of persuasive health applications. They use the dichotomy of *classical* and *expressive* aesthetics, where classical aesthetics are described by words like *simplicity*, *orderliness*, *proportion*, or *symmetry*; while expressive aesthetics are described as *original*, *creative*, *fascinating* or *sophisticated.*

Their main finding was that classical aesthetics were more likely to be perceived as credible, while expressive aesthetics does not influence perceived credibility. The authors suggest that designers of persuasive systems such as fitness applications should focus on these classical aesthetics in order to enhance the credibility of their system.

2.5. Summary and differences to previous research

The literature review showed that there are several opportunities for sleep tracking systems to improve. These opportunities include:

- Current sleep tracking systems do not consider the circadian rhythms of the user, which is an essential aspect of sleep [59, 46] that should be explored further.
- Established design requirements have not yet been implemented. These include *detecting anomalies* in order to provide better feedback [11], *giving feedback in ranges*, *increasing transparency* [43], and providing the user a better understanding of what healthy sleep is and what they can do to better their sleep [28].
- There is a lack of personalised advice in sleep tracking systems, as mentioned by Ravichandran et al. [43], Daskalova et al. [13], and Bauer et al. [4].

Previous work within sleep tracking systems has been either pure user studies or user-tests of already developed prototypes. This project differs in that it uses user studies as part of the prototype development. This method offers different insights into what users want from a sleep tracking systems, that can further shape the resulting prototype. It also offers insights into the development itself, which has not been described in previous work.

3. Methods

This chapter will describe the methods used to gain insight into the use of sleep tracking systems, as well as the development of both low-fidelity and high-fidelity prototypes.

3.1. User studies

Three types of user studies were conducted in order to understand user opinions and behaviours regarding sleep tracking systems. First was an online survey, that aimed to give an overview of the demographic and their opinions on features of sleep trackers. In order to further this insight, a series of interviews were conducted, which focused more on how users integrated sleep tracking into their lives and how current systems are working for them. Based on these insights, wireframe prototypes were produced, which were evaluated using an online usability survey.

3.1.1. Online survey

Surveys are used in HCI research to measure attitudes and intent, as well as user characteristics, technological interactions and awareness. These types of data can enable researchers to identify segments of users, who may need different user experiences [32]. An online survey was used because of its benefits of a wider geographical reach, shorter fielding period, and lower bias due to respondent anonymity [32].

The survey was conducted in order to answer the following research questions: RQ1 –*How do users interact with sleep tracking systems?*–, RQ2 –*What are user motivations for using sleep tracking systems?*– and RQ3 –*What functions do users find necessary in sleep tracking systems?*–.

Survey questionnaire

This section will describe the structure and content of the questionnaire used in the survey. The full questionnaire can be found in Appendix D.

In order to answer how users interacted with sleep tracking systems (RQ1), the questionnaire asked participants if they had used a sleep tracker system in the past. It also asks which kind of sleep tracker they had used, for how long they had used it, how often they look at their sleep data, and what their primary motivation for using a sleep tracker was.

User motivations for using a sleep tracking system (RQ2) were looked into not just by asking what their reason for using a sleep tracker was, but also by asking what other types of tracking activities they engaged in. With the activities diet, activity, and habit tracking, participants were asked to rate their familiarity on a scale from 1 to 5.

One hypothesis was that participants who used devices such as smartwatches would be more likely to be using, or be interested in, sleep tracking systems. Participants were asked whether they owned a smartwatch and their reason for owning one. These participants could choose from one of four possible reasons (notifications, activity tracking, sleep tracking, applications), or they could supply their reasons via a text box.

The questionnaire included fourteen statements that participants answered using a modified *Likert scale* where 1 was *strongly disagree* and 5 was *strongly agree*. Some of these statements were intended to reveal which participants may be receptive to using sleep tracking in the future, while others were intended to gauge what features are essential in sleep tracking systems. The latter also included features that are not present in current systems, intending to identify possible additions to current systems. In order to avoid a *question order bias* these statements were presented in a randomised order.

Similarly, participants were also asked to rate ten features such as *"Alarm function"* and "*Ability to enter additional data*" on a scale from 1 to 5 where 1 was *not important*, and 5 was *very important*. Participants also had the option of adding features via a text field.

For demographic data, participants were asked which age group they belonged to (18-24, 25-34 35-44, 45-55 and 55+) and gender (male, female, other with an optional text field). These sorts of data allow for comparisons between age groups, as well as identifying potential gender differences in the use of sleep tracking systems.

Distribution The survey tool SurveyXact was chosen as it was available through the University of Bergen, and also allowed for anonymous collection of data, which was necessary in order to comply with privacy guidelines. The survey link was shared on social media, including Facebook, Twitter, and Reddit (specifically the *samplesize* community). The sample should be considered self-selected and may not be representative of a general population [27, p. 113-117].

Analysis Statistical analysis was done using the *RStudio* statistical software. First, descriptive statistics were produced to gain an overview of the data. These statistics included the number of responses, the number of answers on close-ended questions, and average ratings on Likert scale items. Differences between groups were illustrated using the *ggplot* package within RStudio. Correlation matrices were used to inspect the relationships between factors visually. Significance tests were used to determine if there were significant differences between groups. The focus of these tests was to determine if there was any significant difference in opinions or viewpoints between user-groups determined by age, use of sleep tracking, smartwatch ownership, or familiarity with similar concepts. Tests were also run on sub-populations such as sleep tracking users, and smartwatch owners, to determine if there were significant differences within these groups.

Two different significance tests were used, depending on the number of groups observed. For comparing between two groups, a Mann–Whitney U test was used. This test was chosen because it does not require the assumption of normal distribution in the dataset. As the sample was self-selected, this assumption could not be made. When there were three or more groups, a Kruskal-Wallis (One-way ANOVA on ranks) test was used. The significance level of 0.05 was chosen in order to avoid Type I errors, as is common practice in HCI research [27, p. 38-39]. In these cases, a Dunn's post hoc test was used in order to determine the difference between the groups.

3.1.2. Interviews

Interviews were conducted with nine users of existing sleep tracking systems. The goal of these interviews was to identify views that users have on current systems and how the current systems are integrated into users lives (RQ2, RQ3). Another benefit is that interviews can give insights into applications that would otherwise require long term use of, and purchase of, these different systems. It also acts as a way to connect with potential users, which is a core tenant of user-centred design [37].

Interview questions A list of prepared questions (see Appendix B) formed the backbone of the interviews, but the interviews were not limited to this list, and as such, it could be categorised as a semi-structured interview. These questions related to their use of the system, including questions about how long they had used the system and why they started using the system. Other questions were about how they used the system: how often do they open the app, at what time during the day, and in which ways they find it helpful. In some cases, questions were specific to the system: Fitbit users were asked about the *Sleep Insights* functionality, as this was a feature that was particularly relevant for this project. Participants were asked if they had made any changes in their day or their day-to-day life because of what their sleep tracking system told them. The goal of this was to identify if and how these systems can affect change in behaviour.

Ethical approval Since the interviews were to be audio-recorded and therefore could be personally identifiable, approval for processing this data was needed. The Norwegian Centre for Research Data, which acts as an advisor to the University on this matter, handled the approval process. After submitting a notification form about the project, which included a data management plan and the information letter that the participants would sign (see Appendix A), approval was eventually granted (as can be seen in

Appendix B).

Recruitment Initially, an attempt was made to recruit local participants by placing posters on the University of Bergen campus. This attempt proved ineffective, as only one participant was recruited using this method.

The criteria for participating were to have used at least one sleep tracking application or device and to be able to give consent to the interview, which means they had to be at least 18 years of age.

After the decision was made to change the interview format to allow for remote interviews, recruitment strategies also changed. Participants were recruited via the social media platform Reddit, which allows users to subscribe to *subreddits* that interest them. Posts were made to the "Quantified self", "Fitbit", "Apple Watch", "Garmin" and 'Sleep' subreddit, as they were related to sleep tracking. The posts linked to a survey where the respondents could enter their email to be contacted about participating. This was done so that the participant could not be tied to their user account. Over 50 people completed this survey, but only ten responded to the follow-up email, with eight ending up participating in the interview.

Procedure Participants who had expressed their interest were contacted via email to schedule an interview. This email also contained the consent form, which they were asked to sign before the interview. The majority of the interviews were conducted via video calls. Participants were asked which form of call service they preferred, and in most cases, this service was used. This was an attempt to lessen the burden of participating, or *meeting users where they are*. The single interview that was conducted in person was conducted at a cafe located at the university campus.

Recording of the interview did not start until verbal consent was given, in addition to the signed consent form. This both acted as another confirmation, and as a formal start to the interview. Notes were also taken during the interview, using pen and paper. These notes were kept simple, using informal codes and timestamps to record relevant statements to return to during the transcription.

Analysis After the interview, an effort was made to transcribe the audio while it was still fresh in the mind of the interviewer. Only the most concise statements were transcribed ad verbum; otherwise, informal codes were used to note the sentiment of the participant.

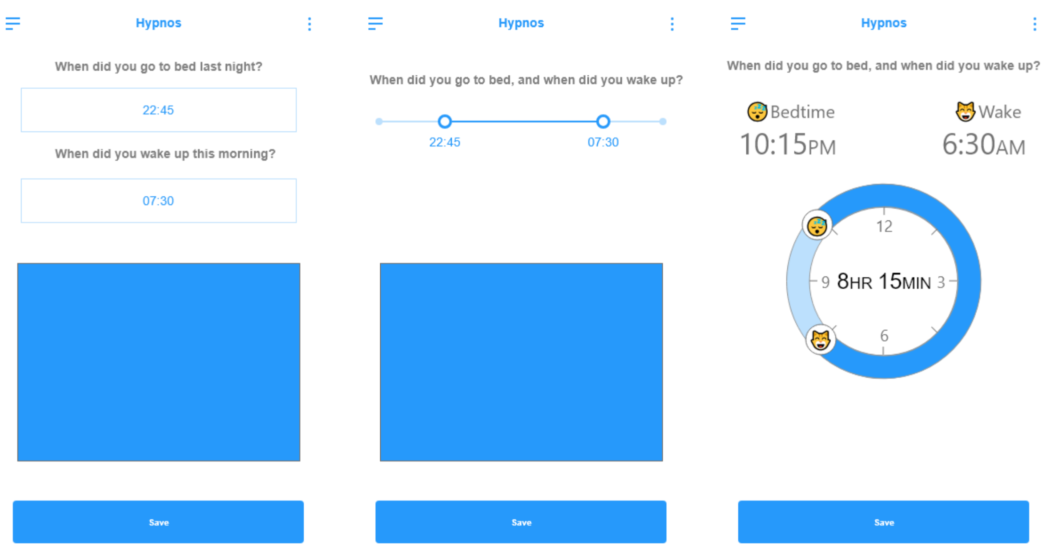
After several interviews were conducted, formal coding of the responses began. Some of these codes were based on previous categorizations (for example [47]), while other codes came out of comparing the interviews as per the *grounded theory method* [33]. As is encouraged with grounded theory, since it can reveal new directions of inquiry, coding was started early in the interview process. This naturally led to a different focus in later interviews than earlier ones, as competing theories started to form. These codes were then categorised, and similarities and differences between categories were further compared. These categories allowed for further identification of what current systems do right, and what they can do better, and what the users wish that they could do.

3.2. Wireframe prototyping

To further understand which features are important to users (RQ3), an evaluation was conducted using wireframe prototypes. Three sets of wireframes were made using Adobe XD, where each set was oriented around the same functionality but used different interactions or information. Care was taken to use a similar graphical style within the set, to avoid it influencing the comparison.

The decision to focus on manual input was made for multiple reasons. It greatly reduced the complexity of the system, making it possible to create a functioning prototype that focused on the user experience. Previous research had also shown that journaling and subjective input was necessary in order to be able to give personalised and actionable advice, which was a goal of this project.

Interface 1: input of bedtime and wake-up time.



(a) (b) (c)

Figure 3.1.: Three wireframes that allows the user to input time slept

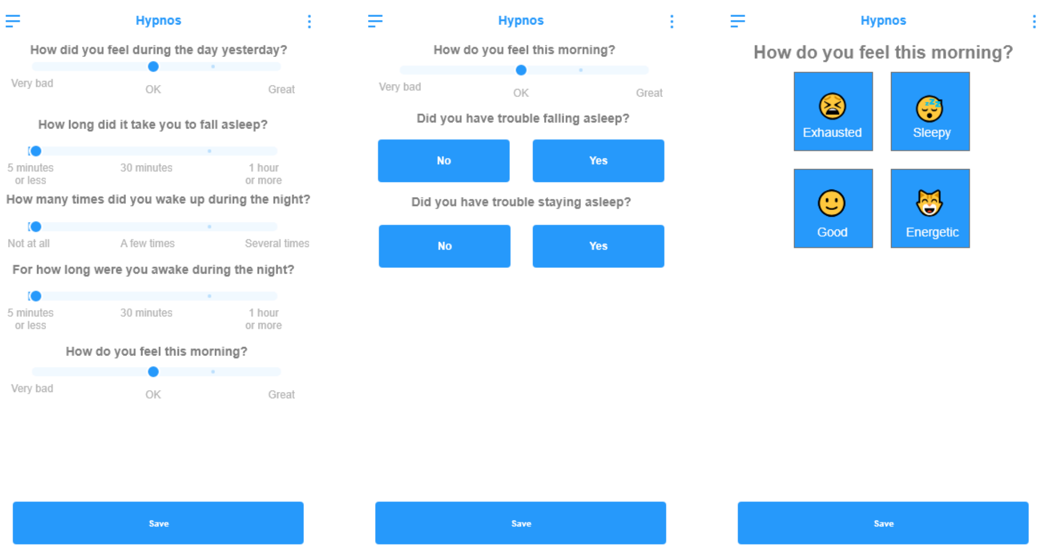
The three wireframe prototypes for *interface 1* can be found in figure 3.1. This interface is meant to allow the user to enter the time they went to bed at night, and when they woke up in the morning. Each prototype uses different input mechanisms, in addition to slight differences in the information presented to the user.

The first wireframe asks the user two questions: the time they went to sleep, and when they woke up. Each question has its own input field underneath. This is meant to resemble how time is often input on the web, and in many smartphone applications. This wireframe leaves it to the user to imagine how the times are inputted, something that could change depending on what operating systems, or devices they are used to.

The second wireframe asks a single question: "When did you go to sleep, and when did you wake up?". Underneath is a multi-handle "range" slider, which is an often-used interaction on the web and mobile phones. In this style of slider, the user can manipulate both the start and end of a range, which in this instance would be a period of sleep. This range is also indicated by each handle having an attached time below it.

The third and final wireframe is somewhat similar to the previous wireframe, but here the range is circular, like a clock. This wireframe is inspired by Apple's "Bedtime" functionality[57] which is included on newer iPhone devices. This wireframe also includes the duration in hours, in addition to the start- and end-point.

Interface 2: Input of subjective sleep quality



(a) (b) (c)

Figure 3.2.: Three wireframes that allows input of subjective sleep quality

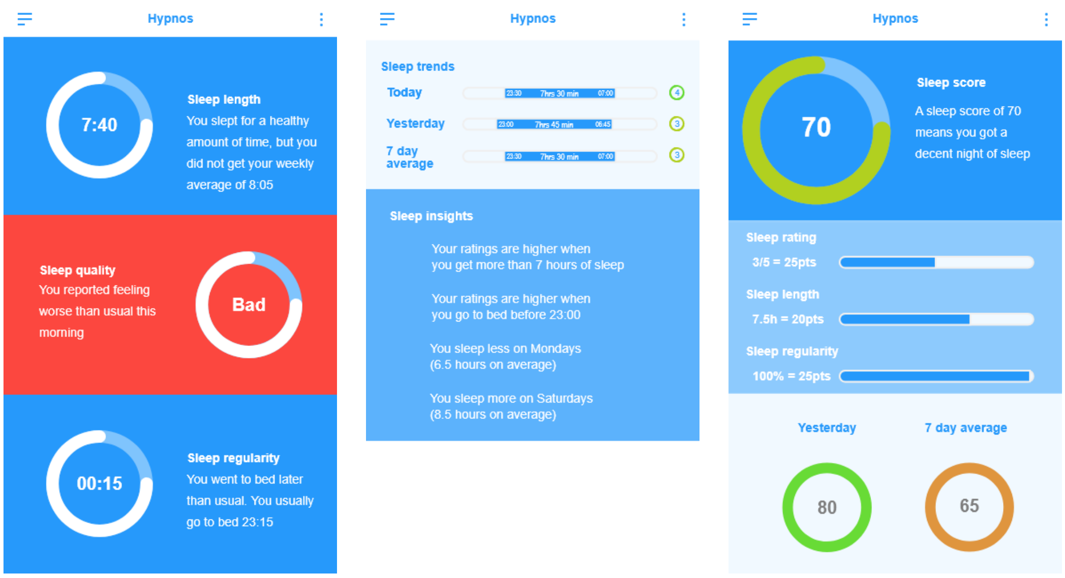
The three wireframes for *interface 2* can be found in figure 4.15. These wireframes are meant to allow users to give their opinion on their sleep quality for the previous night.

The first wireframe uses five sliders to allow the user to input how they felt during the day, how long it took them to fall asleep, how many times they woke up during the night and for how long, how they felt this morning. This interface is modelled after the sleep diaries used in CBT-I, which are further described in chapter 2.

The second wireframe narrows it down to three questions: "how do you feel this morning?", "did you have trouble falling asleep‘" and "did you have trouble staying asleep?". The first question is answerable using a slider, while the latter two use buttons saying "yes" and "no". In many ways, this wireframe can give the same answers as the previous interface (did the user have a restful night, or did they wake up often?), but with less precise data.

The third wireframe asks the question "how do you feel this morning". The question is answerable using four buttons, labelled "Exhausted", "Sleepy", "Good", and "Energetic" that each has an emoji meant to match the label. This interface is inspired by the application Mi Fit, which has a similar interface. Again the complexity for the user is reduced, with the cost of less complex data.

Interface 3: Giving the user feedback on their sleep habits



(a) (b) (c)

Figure 3.3.: Three wireframes that gives the user feedback on their sleep habits

The three wireframes for *interface 3* can be found in figure 4.17. This interface is intended to give the user feedback on their sleep data. The three wireframes differ in both which data they surface to the user, and the methods they use to do so. As with the previous two interfaces, the goal was to limit the graphical differences between the wireframes, but since this was a more complicated interface, there were more visual differences between these wireframes.

The first wireframe gives the user information about their sleep length, sleep quality, and sleep regularity. Each section has an accompanying text that compares the current data to their weekly average or their "usual" sleep quality or regularity. This wireframe uses colour to reinforce this: a blue background is as good or better than usual, while a red background is associated with worse than usual data.

The second wireframe has two main elements. First, there is a "sleep trends" section, where the user can see their bedtime and wake-up time, duration, and rating for the last day, the day before that, and a "7-day average". The second element is "sleep insights". Here the user is presented with text that is supposed to provide insights on their sleep habits. In this wireframe there are four of these texts. Each text shows correlations between hours of sleep and rating, bedtime and rating, as well as which days the user has slept more or less than usual. The intention is for these insights to spur the user to reflect on their sleep habits. In this case, one could imagine the user asking themselves why they sleep less on a Monday, or why they wake up feeling better when they go to bed earlier.

The third wireframe gives the user's most recent night of sleep a score between 0 and 100. In addition to the total score, this wireframe displays the three components of the sleep score: the subjective rating, sleep length and sleep regularity. This wireframe also displays the score of the previous day and a seven-day average score. As with the first wireframe, colour is used to reinforce the score using a gradient from red to green. For the most recent sleep score, the colour of the circle is light green, to indicate that it is good. For the sleep score of "yesterday," the colour is a deeper green, indicating that it is even better, while the "7-day average" is orange, indicating that it is not so good.

Procedure

Each set of wireframes was compared using an online survey, employing the *System Usability Scale* (SUS) questionnaire[7], which is a commonly used questionnaire within usability research. The user was presented with a wireframe and then asked to answer the ten items of the questionnaire. The order of the wireframes was randomised in order to eliminate question order bias. After having evaluated each of the three wireframes, respondents were asked which interface they preferred, how familiar they were with sleep tracking systems, and demographic questions including age, gender and country.

The first survey was distributed via social media (Facebook, Twitter and Reddit) until it reached 25 responses. As this took longer than anticipated, the two following surveys were distributed via Amazon Mechanical Turk instead. This change resulted in a different demographic of respondents between the first survey and the two that followed. As interfaces were not intended to be compared across sets, this was not considered to be an issue.

Analysis

The analysis started with calculating the SUS scores of each wireframe using RStudio, giving a number between 0 and 100. These scores allowed for a comparison between the three wireframes in each set. Wireframes were not compared across sets, as each set had a different interaction context, where users may have different expectations depending on their familiarity or experience. Comparisons were also made between the mean of each of the individual items on the SUS questionnaire. Which wireframe was most preferred was another way of comparing the wireframes. In order to determine if factors such as age, gender, and familiarity with sleep tracking influenced the SUS score, correlation tests were used. This analysis was done in the same manner as with the previous user survey, except that in this case, the Mann–Whitney U test was used exclusively.

3.3. High-fidelity prototype development

As outlined in the media design method [38], designing media can be used as a way to study it and the technology behind it. This method includes building a prototype of a medium, experimenting with different types of content for the medium and evaluating it with users. Evaluations with users in real-world settings are made possible by using high fidelity prototypes, and digital media is making the distribution of such prototypes easier. Another benefit of the high fidelity prototype is that it can uncover what technical aspects of a medium is more challenging to implement since it at least has to simulate the final product. It can also be turned into a product, or form the basis of a product, as long as technical choices do not hinder further development.

The implementation of the prototype is described in chapter 5.

4. Results from user studies

In this chapter, the results from the conducted user studies will be presented. It will present the analysis of the online questionnaire, the interviews, and the wireframe evaluations which were described in the previous chapter.

4.1. Online questionnaire

This section will present the result from the online questionnaire. First, it will give an overview of the responses to this questionnaire using descriptive statistics, and then it will identify significant differences and look at correlations between groups within the responses.

4.1.1. Response overview

Table 4.1.: Overview of respondents

Total	67
Male	43
Female	24
Type of sleep tracker	
Total	31
Smartphone	17
Smartwatch	18
Dedicated	2
Age ranges	
18–24	21
25–34	23
35–44	15
45–54	7
55+	1

In total, the survey had 67 responses, of which 43 were from male respondents, and 24 from female respondents. Responses were mainly from a younger demographic, with a majority of respondents (65%) being in the 18–34 age range. Only one response came from someone above the age of 55.

Sleep tracking use

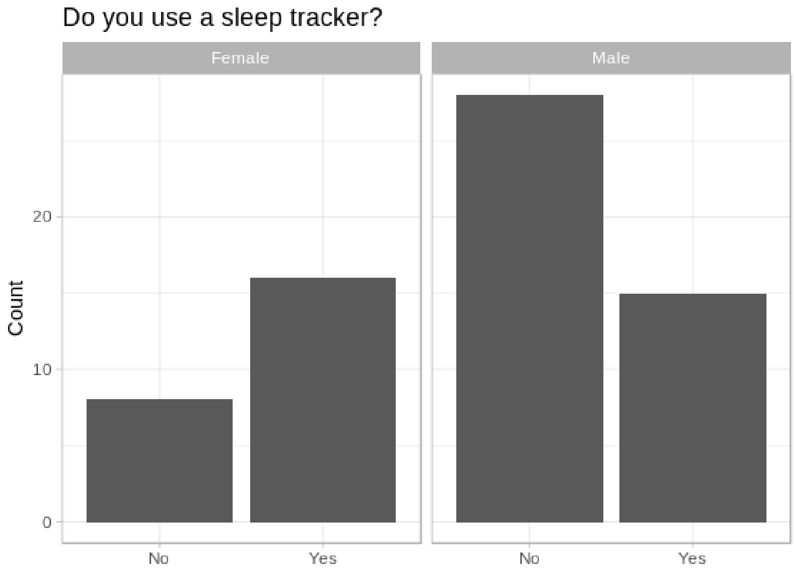


Figure 4.1.: Difference in response by gender

Less than half the respondents (31) reported having used a sleep tracker in the past. Of these, 16 were female, and 15 were male, which show that the women who took this survey were more likely to have used sleep tracking. This is unlikely to be the case in the general population and is likely a result of how the sample was recruited. Of those who had previously used a sleep tracking system, 17 reported doing so on a smartphone, 18 via a smartwatch, and two had used dedicated sleep tracking devices.

Table 4.2.: Reported duration of sleep tracking

Less than 1 month	3
Less than 3 months	8
Less than a year	5
More than a year	15

Close to half of the respondents (15) reported having used a sleep tracker for longer than a year, while a total of 11 had used it for less than three months. Five responded with a period longer than three months, but shorter than a year.

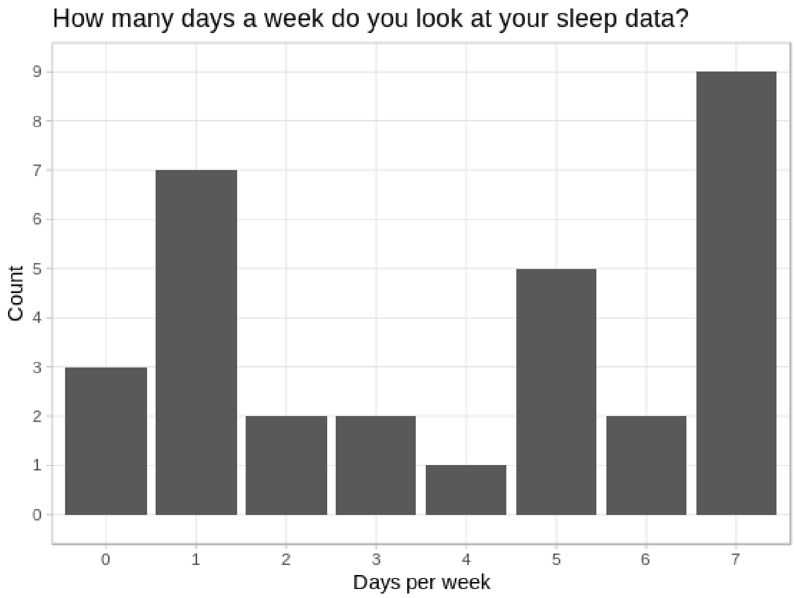


Figure 4.2.: How often respondents interact with their sleep data

When it came to how often the respondents interacted with their sleep data, nine respondents reported doing so every day, and seven did so 5–6 days a week. On the other end of the scale, three respondents did not check their sleep data, while a total of nine checked their data once or twice a week. This may indicate that there are some users who like to interact with their sleep tracker on a day–to–day basis, while others may want to look at it once a week to review a longer period. It may also be a result of how the application presents the data; some applications may be better at delivering day-to-day information, while others are better at showing trends over time.

Sleep tracking devices

In figure 4.3, one can see the distribution of smartphone and smartwatch sleep tracking by age group. When looking at the youngest demographic (18-24), the majority are using smartphone sleep tracking, as opposed to the following demographics where other types of tracking are more popular. There may be several explanations to why this is the case. One could be economical, as younger people often have less disposable income and therefore may not be willing to purchase a device specifically for sleep tracking. It could also be the *ease of access* of downloading an app instead of purchasing a physical device. More speculatively, it could also be a result of this age group being more likely to be a later chronotype, who may have more of a problem getting up at socially acceptable hours. They may also be more interested in the "smart alarm" applications that are more commonly available on smartphones, but not on smartwatches.

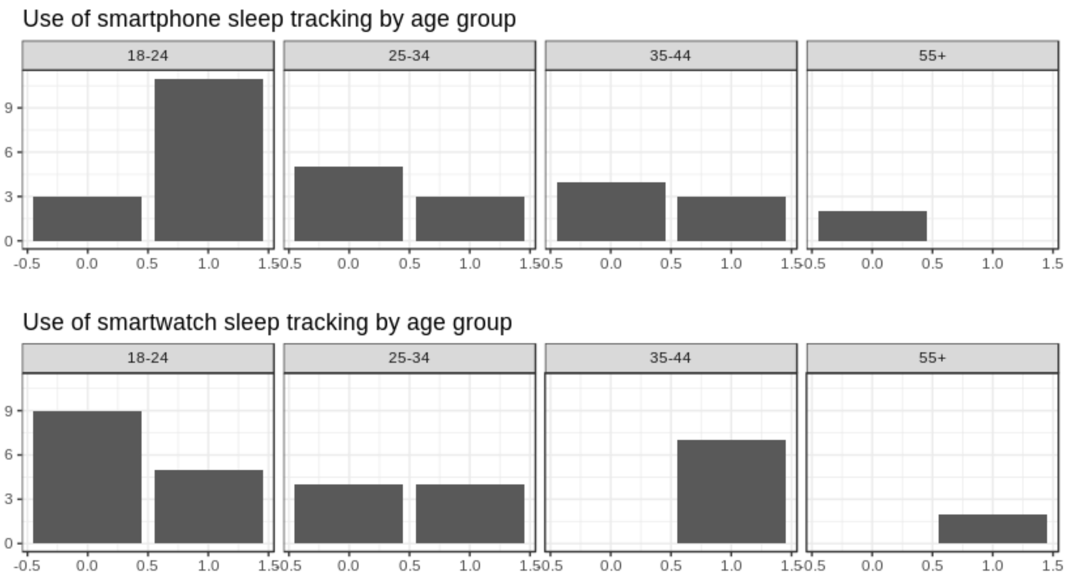


Figure 4.3.: Use of sleep tracking device by age group

Reasons for sleep tracking

Table 4.3.: Reported reasons for sleep tracking

Understand sleep behaviour	24
Alarm function	12
Bedtime reminder	4
Other	6

When asked what their main reasons for sleep tracking were (with multiple-choice allowed), 12 selected "Understand sleep behaviour, another 12 selected "Alarm function", and four selected "Bedtime reminder". Six participants gave other reasons specified via text, which included recording of sleep talking and snoring.

Other tracking

In table 4.4, the other types of tracking that was reported is listed. The most common was activity tracking, followed by financial tracking and diet tracking. Health tracking was also somewhat popular, with mood and habit tracking being less so. *Other tracking*, which included fertility tracking, and specific health-related tracking, was also reported.

Table 4.4.: Types of tracking reported

Sleep	31
Diet	16
Activity	36
Mood	8
Finance	29
Habits	6
Health	13
Other	5

Table 4.5.: Smartwatch owners

Total	23
Female	10
Male	13
Age	
18–24	5
25–34	8
35–44	8
45–55	2

Smartwatch use

Table 4.5 gives an overview of the participants who indicated that they owned a smartwatch. Out of a total of 67 responses, 23 reported using a smartwatch. By demographic, there are more respondents in the 25-34 and 35-44 age groups that reported wearing a smartwatch. As previously stated, the youngest age group were more likely to use a smartphone for sleep tracking, and one can see that they are also less likely to own a smartwatch.

In table 4.6, the reasons given for using a smartwatch is listed. Most common is measuring activity, which almost all listed as a reason. Sleep tracking was also reasonably common, as was receiving notifications. Applications were not as common, which may indicate that this is not the main draw of smartwatches. Another explanation could be that smartwatches were loosely defined in the survey and may include devices that do not run user-installable applications.

Table 4.6.: Reasons given for owning a smartwatch

Notifications	14
Measure activity	20
Sleep tracking	15
Apps	8
Other	2

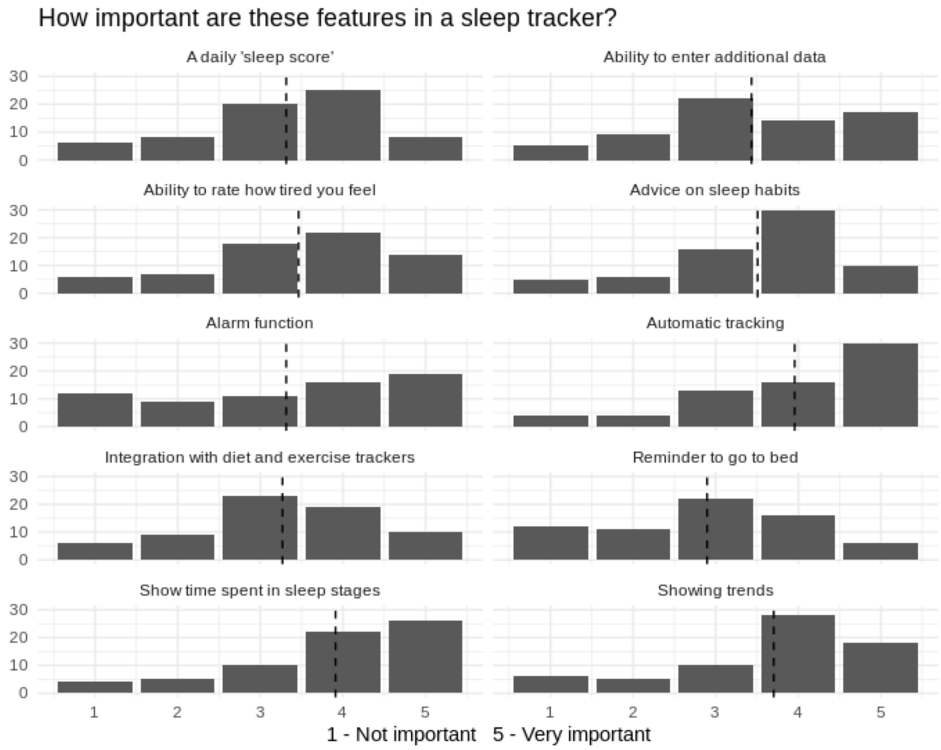


Figure 4.4.: Importance of features

4.1.2. Importance of features

Figure 4.4 shows the participants' responses when asked to rate the importance of different functions of sleep tracking. Not surprisingly "automatic tracking", and "showing time spent in sleep stages" had the highest averages. "Showing trends" also had a high average score.

In figure 4.5 the averages are sorted and grouped by sleep tracking use. Interestingly, there are a few differences between the two groups. "Ability to enter additional data" rated much better with users of sleep tracking, as did "alarm function". On the other hand, "advice on sleep habits" and "a daily sleep score" did better with those who are not tracking sleep. It may be that users of sleep tracking systems feel they know enough

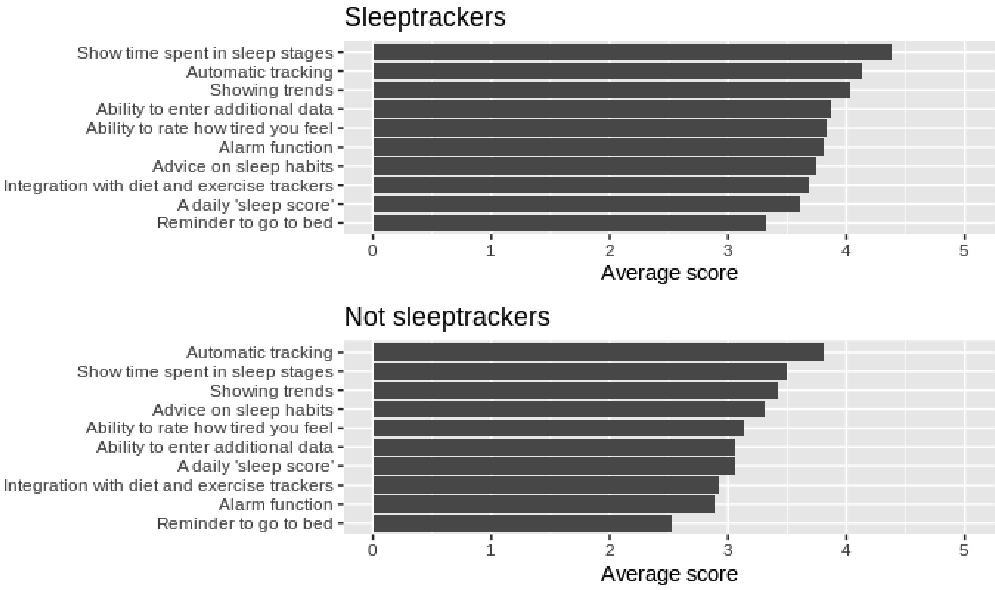


Figure 4.5.: Mean response to features by sleep tracking use

about sleep and do not need these features.

In figure 4.6, one can see the responses to several statements regarding sleep tracking systems. Here 1 is *strongly disagree* while 5 is *strongly agree*. The population is firmly in favour of a "sleeping tracker should show how long I have slept", which is not unexpected. Estimating the quality of sleep also did well. Contrary to the rating of features, the statement "*a sleep tracker should give me a score [...]*" did somewhat better in this context. The same can be said of "I want advice on how to sleep better", which has a different distribution to "advice on sleep habits" in the previous case. Otherwise, the general population seem somewhat positive to statements regarding connecting various activities with sleep. On the statements "*a sleep tracker would help me sleep better*" and "*I plan to use sleep a tracker in the future*", the population is more evenly split, with the latter seemingly having a more extreme division.

When grouping the responses by sleep tracking use as in figure 4.7, the most striking difference is the average score of "I plan to use a sleep tracker in the future" being much lower in the non-sleep-tracking population. One can also see that "*a sleep tracker would help me sleep better*" did poorly even with the sleep tracking population. "*I want advice on how to sleep better*" did better with the sleep tracking population, which is the opposite of how "*advice on sleep habits*" did in the previous set.

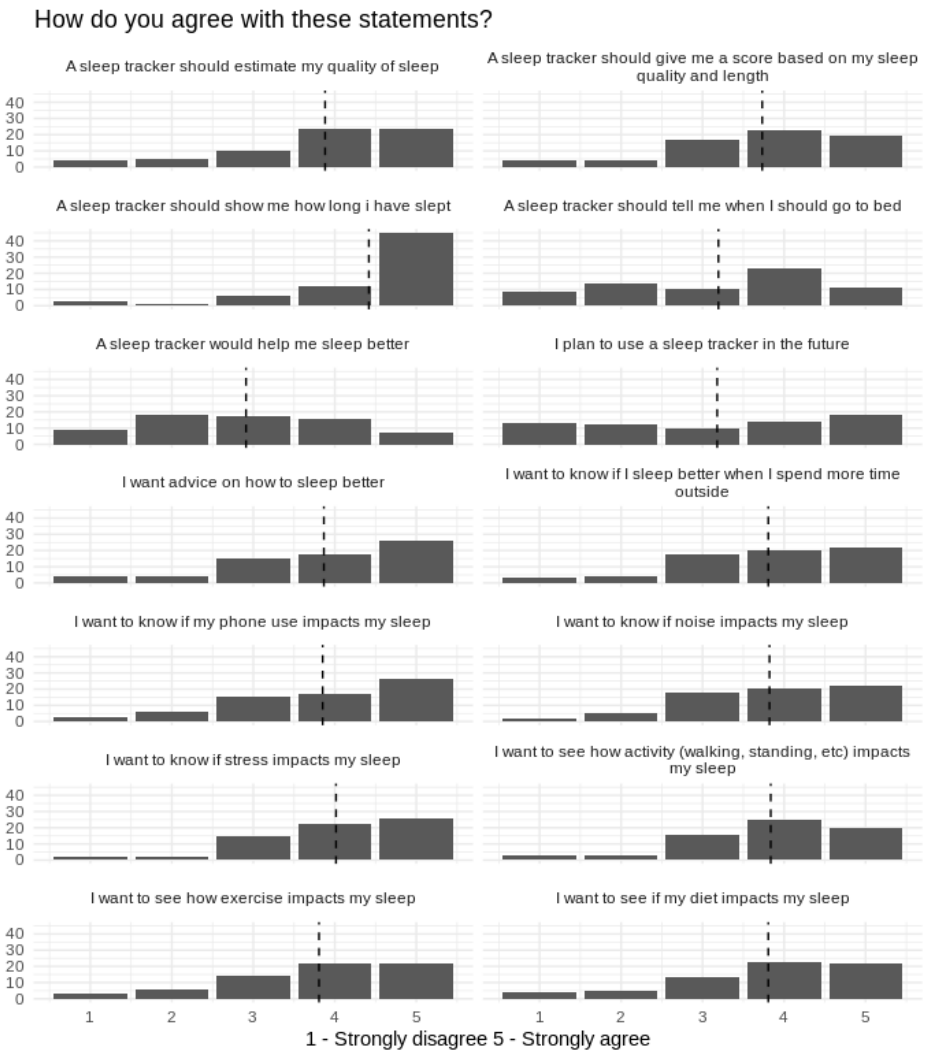


Figure 4.6.: Responses to statements regarding what a sleep tracking system should be

4.1.3. Correlation analysis

Figure 4.8 shows a *correlation matrix* that includes the Pearson correlation between demographic data and responses. Starting with demographic factors, one can see that there is a weak negative correlation between age and use of sleep tracking (r = -0.24), and also between age and tracking finance and habits (r = -0.2 in both cases). With gender, we also only see weak correlations: with financial tracking (r = 0.21) and familiarity with diet tracking (r = -0.32).

Use of sleep tracking is moderately correlated with smartwatch ownership (r = 0.46), and weakly correlated with tracking exercise (r = 0.38), diet (r = 0.25) and familiarity

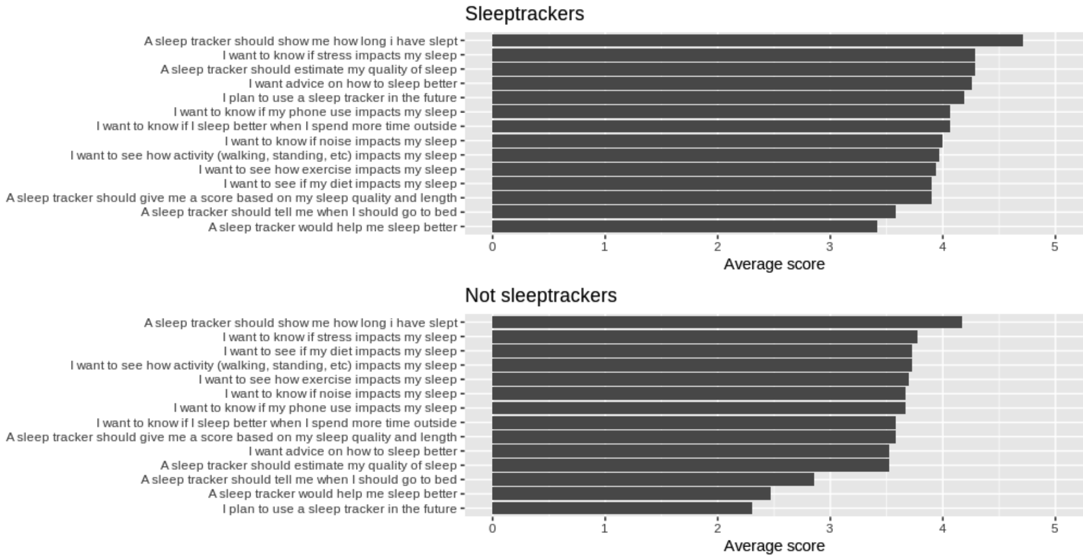


Figure 4.7.: Mean response to statements by sleep tracking use

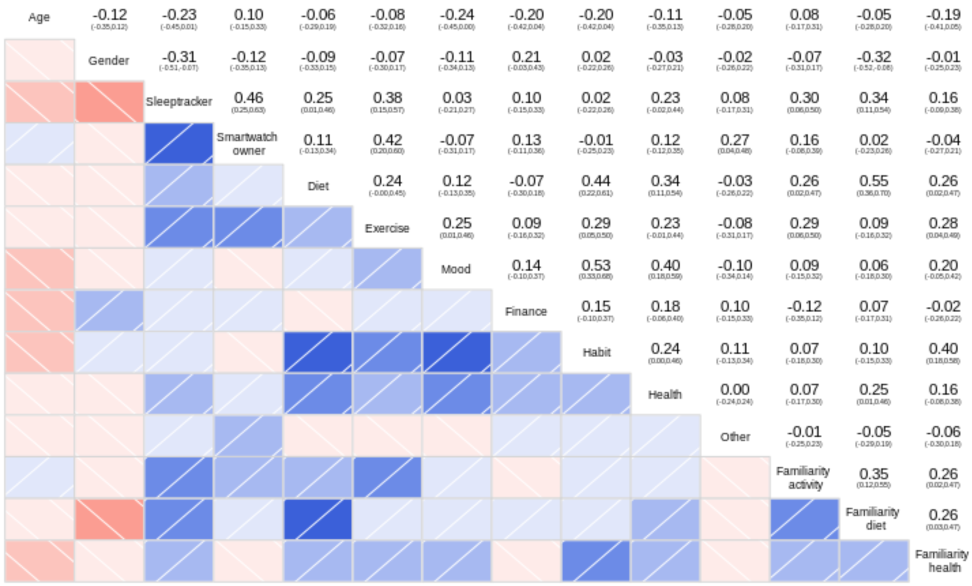


Figure 4.8.: Correlation matrix showing the Pearson correlations between demographic data and responses

with activity tracking (r = 0.3). Smartwatch use is also moderately correlated with exercise tracking (r = 0.42). Tracking of diet is moderately correlated with tracking habits (r = 0.44) and with the familiarity of diet tracking (r = 0.55). Mood tracking also has a moderate relationship with habit tracking (r = 0.53) and health tracking (r = 0.4), suggesting that these forms of tracking appeal to the same user group. That habit tracking has only a weak correlation with health tracking (r = 0.24), but a more moderate correlation towards *familiarity* of health tracking (r = 0.4), may indicate that this user group is aware of this type of tracking but does not engage in it. Health tracking has a weak correlation with the familiarity of diet tracking (r = 0.25).

Between familiarity with the different types of tracking, there are only weak correlations ($r < 0.35$), suggesting that familiarity with one kind of tracking does not predict familiarity with another.

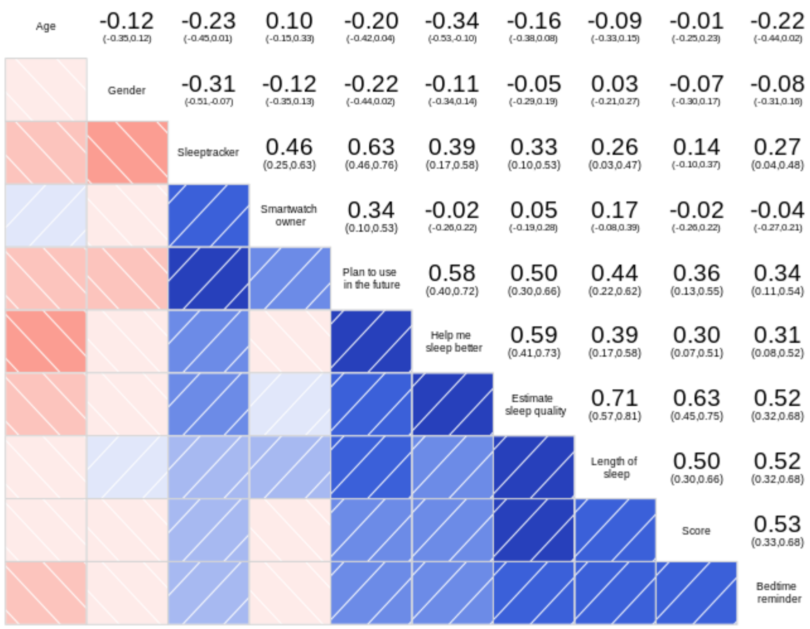


Figure 4.9.: Correlation matrix showing the Pearson correlations between demographic data and responses

In figure 4.9, one can see correlations between responses to questions about what a sleep tracker should be able to do. With age, there is a weak negative correlation with "*a sleep tracker would help me sleep better*" (r = -0.34). With gender, we only see a weak negative correlation towards planning to use a sleep tracker in the future (r = -0.22). Use of sleep tracking is somewhat strongly correlated with planning to use sleep tracking in the future (r = 0.63) and weakly with believing "*a sleep tracker would help me sleep* *better*" (r = 0.39). We also see weak correlations with "*estimating sleep quality*" (r = 0.33), "*length of sleep*" (r = 0.26) and "*bedtime reminder*" (r = 0.27).

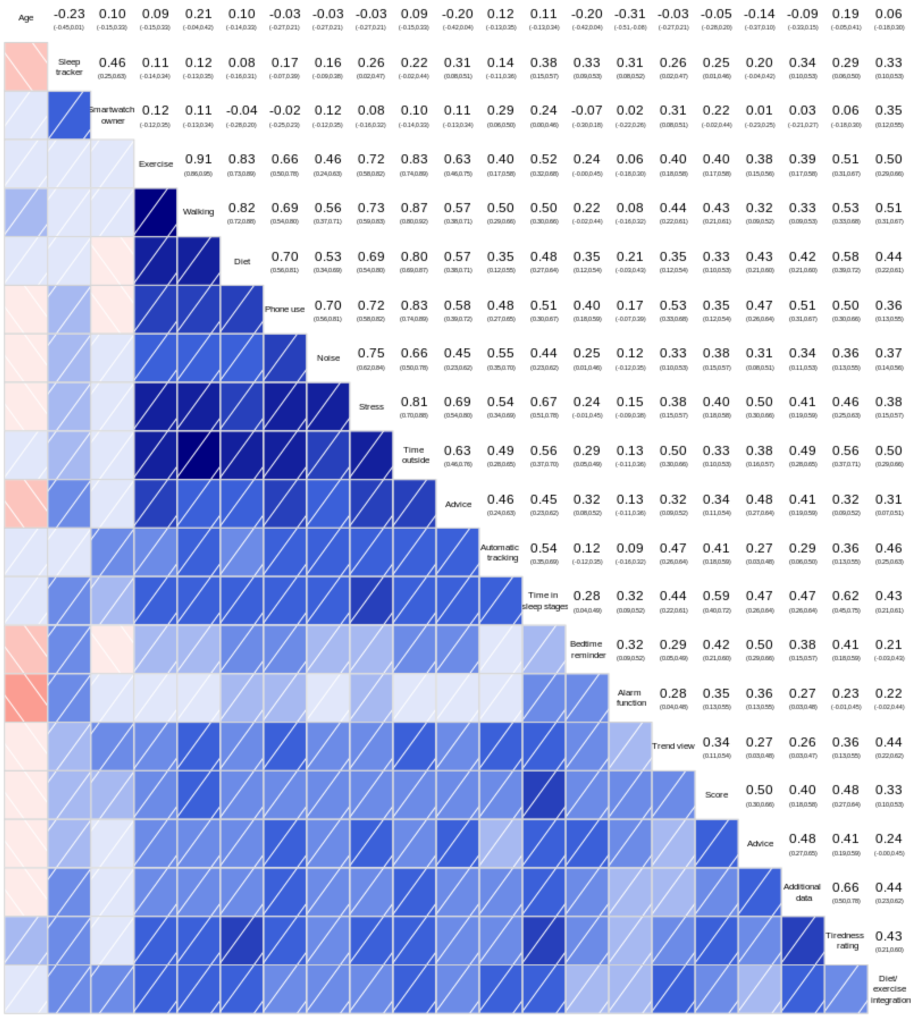


Figure 4.10.: Correlation matrix showing the Pearson correlations between demographic data and responses

Smartwatch use is also weakly correlated with planning to use a sleep tracker in the future (r = 0.34). Planning to use sleep tracking in the future is moderately correlated with believing it will help them sleep better (r = 0.58), that it should estimate sleep quality (r = 0.5), display length of sleep (r = 0.44). Towards the sleep tracker giving a score, we see a weak correlation (r = 0.36).

Believing that a sleep tracker will help with sleep is moderately correlated with estimates of sleep quality (r = 0.59) and weakly with displaying the length of sleep (r = 0.39). There are strong correlations between estimating sleep quality and length of sleep (r = 0.71) and with a sleep score (r = 0.63). Between the importance of sleep quality estimation and a bedtime reminder, there is a moderate correlation as well (r = 0.53). The importance of the length of sleep is moderately correlated with displaying a sleep score (r = 0.5) and with bedtime reminder (r = 0.53). Between the sleep score and bedtime reminder, there is also a moderate correlation (r = 0.53).

4.1.4. Significance testing

This section describes the results of significance testing, as described in the previous chapter. It will begin by looking at differences between groups based on gender and age, before looking at differences between groups based on sleep tracking use and smartwatch use. Finally, it will describe the significant differences within the subpopulation of sleep tracking users. Here, responses are also grouped by length of use, and which device they use.

Gender and age Significance tests related to gender did not provide any new insights. It showed differences in sleep tracking use between the genders but, as previously mentioned, this is likely affected by the sampling method and therefore not transferable to the general population. Otherwise, it confirms that females are significantly more familiar with diet tracking (p = 0.015), within this population. When looking at differences between age groups, testing found that younger sleep trackers were more favourable to the statement "*A sleep tracker would help me sleep better*" (p = 0.018).

Sleep tracking use Those who track their sleep were more familiar with other types of tracking. This includes diet (p = 0.04), exercise (p = 0.002). There is also a significant difference in smartwatch ownership (p = 0.0001), and planning to use a sleep tracker in the future (p < 0.0001). Significant differences are also found in the belief that sleep tracking can improve sleep (p = 0.0013), if a sleep tracker should estimate the quality of sleep (p = 0.0122) and if a sleep tracker should tell the user to go to bed (p = 0.0031).

In terms of interests in features, significant differences were found in the following statements when grouped by use of sleep tracking:

- Stress impact on sleep (p = 0.0461)
- Advice on how to sleep better (p = 0.0214)
- Showing time spent in sleep stages (p = 0.0036)
- Bedtime reminder (p = 0.0106)
- Alarm function (p = 0.0156)
- Sleep score (p = 0.038)
- Ability to enter additional data (p = 0.0086)

- Ability to rate how tired one is (p = 0.0308)
- Diet and exercise integration (p = 0.0086)

Smartwatch use With owners of smartwatches there are significant differences in use of sleep tracking ($p < 0.001$), tracking exercise ($p < 0.001$) and tracking "other"(p = 0.027). There is also a difference in planning to use a sleep tracker in the future (p = 0.0053). The significant differences in responses to interests in features includes:

- Wanting automatic tracking (p = 0.0338)
- Sleep stage tracking (p = 0.046)
- Trends (p = 0.0165)
- Sleep score (p = 0.0406)
- Diet and exercise integration (p = 0.0056)

Subpopulation: sleep tracking users When the data-set was limited to sleep tracking users only (n = 31), significance tests revealed additional differences.

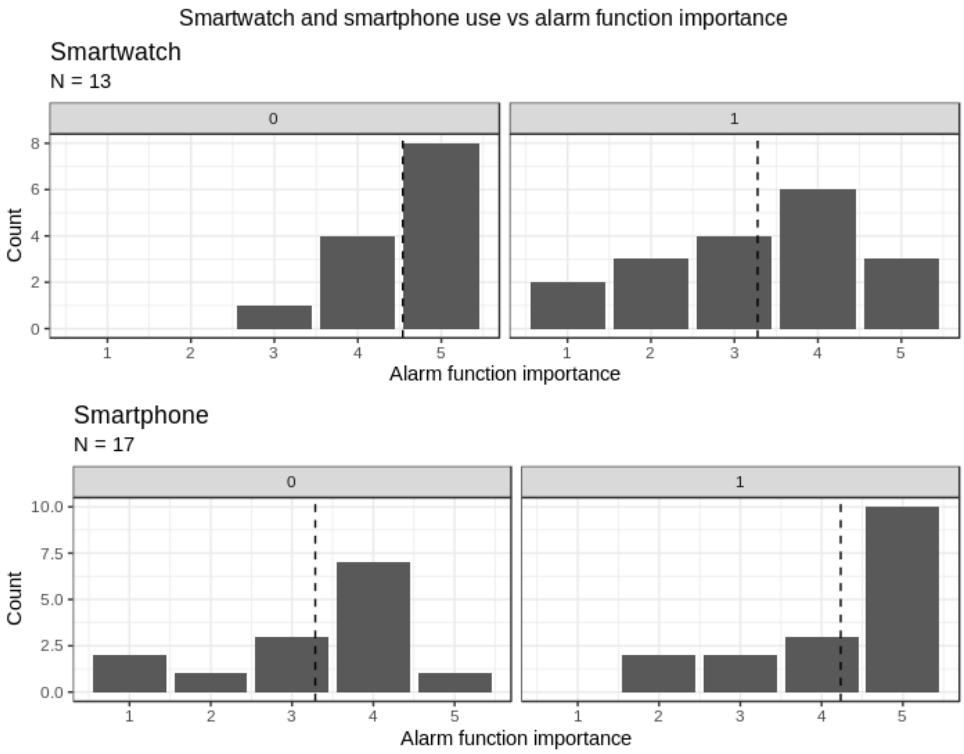


Figure 4.11.: Differences in alarm function importance by device

When grouped by age, there is a significant difference in the importance of alarm function (p = 0.0107). When grouped by length of sleep tracking, there is a difference in familiarity with activity tracking (p = 0.0464). There is a difference in importance of alarm function (p = 0.0153) when grouped by use of smartphone-based sleep tracking systems, which is further illustrated in 4.11. When grouped by smartwatch use, there are significant differences in the importance of automatic tracking (p = 0.0139) and alarm function (p = 0.0034).

4.1.5. Summary

- Many interact with their sleep systems every day, or close to it. There is also those who interact on a weekly basis. This means that sleep tracking systems should support both daily use, perhaps in the form of daily reports, but also provide a good experience for those who prefer to use it less often.
- A majority responded with wanting to understand their sleep behaviour as their main reason for using a sleep tracking system. Some also gave an alarm function as a reason. The latter is most likely using a smartphone for sleep tracking, as this is not a feature of most other devices.
- Younger respondents (age 18-24) were more likely to be using a smartphone app to track their sleep, while later age groups are more likely to use a smartwatch to track their sleep. This could mean that a new sleep tracking system that is targeted at the younger demographic should be smartphone-based. It also means that such a system targeted at the older demographics should have some form of integration with smartwatches or be smartwatch based.
- When asked which other types of tracking they engaged in, activity tracking was the most common (36), with financial tracking (29) and diet tracking (16) being other common types of tracking. As both diet and activity can affect sleep, a sleep tracking system looking to integrate other kinds of tracking should focus on these. Doing the same with financial tracking may be more difficult, but it is also a possibility.
- Of the respondents who owned a smartwatch (23), 15 responded with sleep tracking being a reason for wearing a smartwatch.
- Users of sleep tracking systems in the youngest age group were significantly more positive when asked if a sleep tracking system would help them sleep (p = 0.018). This is something that should be further investigated, as it could mean several things. One explanation could be that this age group are lacking in education on sleep and therefore are more interested in a sleep tracking system. It could also be that they struggle more with sleep, as they could have a later chronotype (see chapter 2). A sleep tracking system could attempt to solve this by providing information on healthy sleep.
- Users of sleep tracking systems are more familiar with other kinds of tracking in general and are more positive towards the benefits of sleep tracking. They are also more positive to features such as a sleep score, alarm function, bedtime reminder, adding additional data, and diet and exercise integration.
- Smartwatch users are more interested in automatic tracking and sleep stage tracking. A sleep tracking system aimed at this audience should make sure to integrate

these functions.

4.2. Interviews

The following section presents the findings from interviews with nine users of sleep tracking systems. The systems used included Fitbit (5), Oura (1), Garmin (1), Apple Watch (1), and a self-made system (1). Most participants had used a sleep tracking system for at least a year, but there were also some who had used it for less than a year. Several users also had experience with systems besides the one they were using at the time.

Table 4.7.: Interview subjects

ID	Gender	Age	Country	Device	Duration
P1	male	32	United Kingdom	DIY	less than a year
P2	male	34	Austria	Oura ring	1.5 years
P3	male	28	United States	Fitbit Charge HR	3 years
P4	female	33	United States	Fitbit Versa	5 years
P5	male	29	United States	Fitbit Ionic	6 years
P6	male	29	Ghana	Fitbit Versa 2	1 month
P7	male	36	Australia	Apple watch (Sleep watch)	3 months
P8	male	22	United Kingdom	Garmin	1 year
P9	male	21	Norway	Fitbit	2 years

4.2.1. Types of tracking

Each participant's relationship with sleep tracking was categorised using the categories of tracking suggested by Rooksby et al. [47]. These categories include *goal-oriented*, *documentary*, and *diagnostic* tracking.

Most participants could be considered goal-oriented in the sense that they want to improve their fitness or sleep. Seven participants can be considered goal-oriented trackers. Their goals range from broader fitness goals, such as improving in sports (P2), to more sleep specific goals such as getting a set amount of sleep (P3, P8, P9), and getting a certain quality of sleep (P1, P6, P7).

Some participants could also be categorised as documentary trackers since they are interested in gathering a set of data that they can analyse at a later period (P3, P5, P8, P9). Worth noting is P5, who does not have a set goal for his tracking, and has some altruistic motives: "I hope that making this [data] available to researchers will allow us to improve these things in the future, and get insights into how these things affect health".

P1 and P4 could be considered diagnostic trackers. P1 aims to reduce his snoring and has experimented with several methods for this. P4 has medical conditions that are causing her poor sleep, and sleep tracking is helping her figure out how treatments are affecting her.

4.2.2. Frequency and time of day

The majority reported checking their sleep data every day, and many did so in the morning. Only P3 and P5 reported looking at it less often. P3 does so weekly or if he experiences poor sleep, while P5 seeks out sleep data a few times a week.

Keeping in mind that those who are willing to participate in an interview may be more enthusiastic about the topic, this result lines up with the findings from the previous survey as it shows that there are users who interact with their sleep tracking system daily. Those who interacted less often seems to have an inclination towards longer-term analysis, one being a programmer, and the other a student of statistics.

4.2.3. Sleep data uses

When asked if their sleep tracking had impacted their lives in any way, several participants responded that it had. Three participants (P1, P2 and P7) reported having made changes during the day after looking at their data. P1 reported working out more after a bad night of sleep to tire himself out for the next night. P2 uses his Oura ring's sleep score to gauge what kind of work he should focus on, and the "readiness" feature to determine how hard he should work out. P7 reported taking naps after seeing he had slept poorly and had also increased the intensity of his workouts after a good night of sleep.

Six of the participants had made changes based on the collection of data and correlating it with other kinds of data. P1 reported having seen an improvement in their sleep after increasing physical activity. P2 had experimented with keeping a tighter bedtime schedule and managed to improve his sleep efficiency. P4 tried different medications based on her sleep data. P6 used sleep tracking in an effort to increase his sleep duration and has found positive improvements in resting heart rate. P7 had noticed that alcohol reduces his sleep quality and is therefore motivated to reduce his intake. P8 had made an effort to reduce screen time before bedtime, and that found it had an impact.

4.2.4. Perceived accuracy of sleep tracking systems

Several of the participants had questions about the accuracy of the trackers. One participant (P1) had previously used a Fitbit for sleep tracking but ended up creating his own sleep tracking system because he did not feel the Fitbit's data lined up with how he felt after sleeping. His solution measures sound, which may fit better with his goal of reducing snoring. Other participants also had felt a disconnect between how refreshed they felt after sleeping and what their systems told them.

P2 had also tried other systems in the past, including Apple Watch and a Garmin, before deciding on the Oura ring. The "sleep score" functionality of the Oura seemed to be the deciding factor for him. P4 had a peculiar problem when she tried a "smart" mattress pad: her cats made it register sleep during the day, which threw off the data. When it came to her Fitbit, she was unsure if it registered the right number of awakenings during the night, which was something she was particularly bothered with. P6 wished his Fitbit could give him accurate data about sleep stages when he slept for less than 3 hours, which was not uncommon for him. P8 had issues with his Garmin registering sleep too early if he was reading or watching TV in the evening. P9 was unsure if the sensitivity of his Fitbit was correct: it showed more awakenings than he remembers.

As five out of nine participants had questions about the accuracy of their systems, one can conclude that sleep tracking systems can improve in this regard. To some degree this can be the result of inaccuracies in the tracking technology itself, but how they present the data to the user could also influence how the system as a whole is perceived. Several users mentioned that their data did not agree with how they felt. If these systems are more accurate than users give them credit for, the creators need to convince the users of this. This could include exploring different methods of communicating, to make sure the user understand meaning of the data.

4.2.5. Other tracking

All of the participants reported tracking things besides sleep. The most common was activity tracking and exercise tracking. This is not unexpected since most of the participants used devices like Fitbit and Garmin that also tracks steps and heart rate. P1 used a Fitbit for activity tracking, but not for sleep tracking. P2 used his Oura for sleep tracking, but another device for activity tracking, even though the Oura has activity tracking functionality. He also reported tracking diet, among other things. P3 also uses his Fitbit for running, and P5 did so for activities like snowboarding and hiking. P6 reported using several functions of the Fitbit app, including logging food, water intake, heart rate, activity and exercise. P7 used another device for tracking heart rate during exercise (MyZone). The application Sleep Watch also allowed for tracking of factors that may influence sleep like alcohol intake and fatigue. P8 also used his Garmin for activity and exercise tracking. He was also participating in Virgin Pulse via his workplace, which allowed for social sharing and competition. P9 used a different device (Polar) for tracking exercise, even though his Fitbit had similar functionality because he preferred the user experience.

4.2.6. Advice from sleep tracking systems

Some of the systems used attempts to give advice based on the data. One example is Fitbit's *Sleep Insights*, which supposedly can "[...]help you learn more about your sleep stats and how your trends compare to those of the same age and gender."[52]. When participants using Fitbit was asked about this feature, several found it to be not personalised enough (P3, P5, P9).

When P5 was asked if he remembered any useful advice: "Nothing comes to mind [...], it's all pretty basic stuff like 'don't exercise too soon before bed' and that kind of thing. I guess it is good to be reminded, but I've never seen one where I go: 'Woah, I did not know that'". P4 was aware of the feature, but did not find it useful: "I have seen it, but I actually have not ever looked at it".

The newest Fitbit user, P6, was the most positive to this feature: "[S]ometimes when I sleep late it tells me to sleep earlier". He also had another example of the system acting smart: "One day I took alcohol before sleeping, and when I woke up the sleep insight told me that when I take alcohol, it is going to affect my REM sleep".

P8 had found the sleep insights useful when he previously used a Fitbit and found that he wanted his Garmin to give similar advice. He noted that Garmin has a function for recommending recovery after exercise, and thought they could do it similarly.

P2, who uses the Oura ring, has found benefit in the sleep score the application gives him: "Since I have been using this for [...] 20 months, I kind of know-how that score reflects my coming day. So if I get a score below 70, for example, I know 'Ok, this is not going to be a great day for any intellectual work'[...]".

P7 had both positive and negative remarks about the advice given by Sleep Watch. He enjoyed that it gave advice based on activity level in relation to sleep, but disliked that it used BMI to compare him to other users. Perhaps a problem particular to fitness trackers is that some users may be too fit.

4.2.7. Summary

Most of the users interviewed interacted with their sleep tracking system on a daily basis. The common reasoning seems to be that they want an explanation for how they feel in the morning. They often found a reason for why they had been sleeping poorly or well. Some had questions around accuracy, which could be an indication that the sleep tracking systems in question should provide better explanations for, or increase the transparency, of their data.

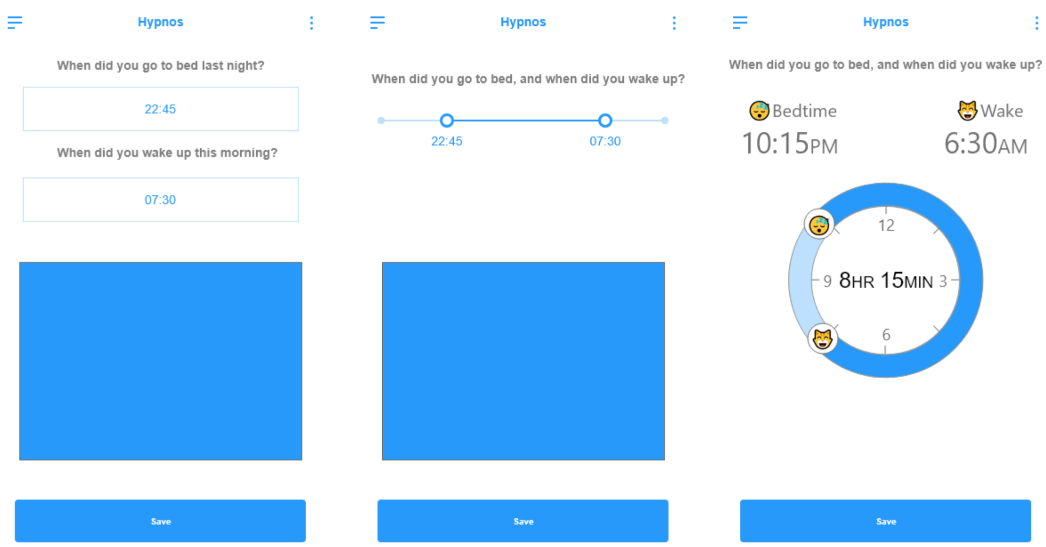
Those that interacted less often were usually more focused on trends. These users often preferred views that allowed them to look at several days, or even weeks, at a time. Some did experiments, where they tried to change something else in their lives, such as exercise and used sleep tracking as one measure of results. Most of those who were interviewed tracked other things. Activity tracking was the most common, as many used a wearable device. This could mean that integration with other tracking is potentially an important feature, either by direct integration or by making it easy for the user to add supplemental data.

Many wanted more personalised advice than their current systems provide. Few of those who used Fitbit devices were impressed by Fitbit's *sleep insights*, which is the only system in this series of interviews that gave advice based on sleep data. They seemed to want advice that was more specific to them, instead of general advice that fits their situation.

4.3. Usability testing of wireframe prototypes

Following the survey and interviews, several designs were made for potential interfaces. These interfaces were eventually pared down to three: one that allowed the user to enter time slept, another focused on input of subjective data, and one that relays information back to the user. For each category, three wireframe prototypes were made, which was compared using a SUS questionnaire.

4.3.1. Interface 1: input of bedtime and wake-up time.



(a) (b) (c)

Figure 4.12.: Three wireframes that allows the user to input time slept

Usability score

Table 4.8.: Mean responses to SUS questionnaire and SUS score

Set 1	"clock"	"fields"	"range"
I think that I would like to use this interface frequently	3.56	3.33	3.07
I found the interface unnecessarily complex	2.3	2	2.26
I thought the interface was easy to use	3.78	3.74	3.59
I think that I would need the support of a technical person to be able to use this interface	1.7	1.96	1.48
I found the various functions in this interface were well integrated	3.44	3.52	3.37
I thought there was too much inconsistency in this interface	2	1.85	2.07
I would imagine that most people would learn to use this interface very quickly	3.96	3.74	3.52
I found the interface very cumbersome to use	2	1.89	2.37
I felt very confident using the interface	3.93	3.85	3.67
I needed to learn a lot of things before I could get going with this interface	2	2.11	1.96
SUS	70.9	71.6	67.6
Percentage of participants preferred	55	30	15

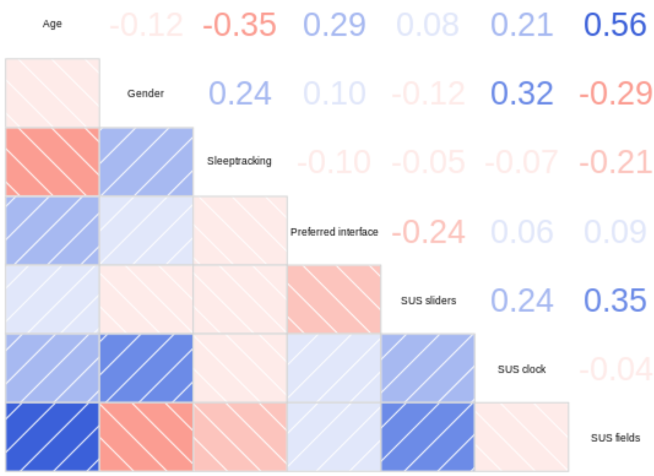


Figure 4.13.: Matrix showing the correlations between different factors and the SUS scores

The three prototypes had a similar usability score, with "clock" and "fields" scoring over 70. While the "fields" interface scored slightly better in total, more respondents—55%—selected the "clock" interface when asked which interface they preferred. Comparing the mean scores of the two, one can see that the "clock" interface scores better in several categories, including frequency of use, ease of use, the confidence of use, and learnability. On the other hand, it scored worse in complexity, inconsistency and integration. This may be a result of the type of prototype, as a non-interactive prototype requires the participant to imagine the complexity themselves instead of experiencing it.

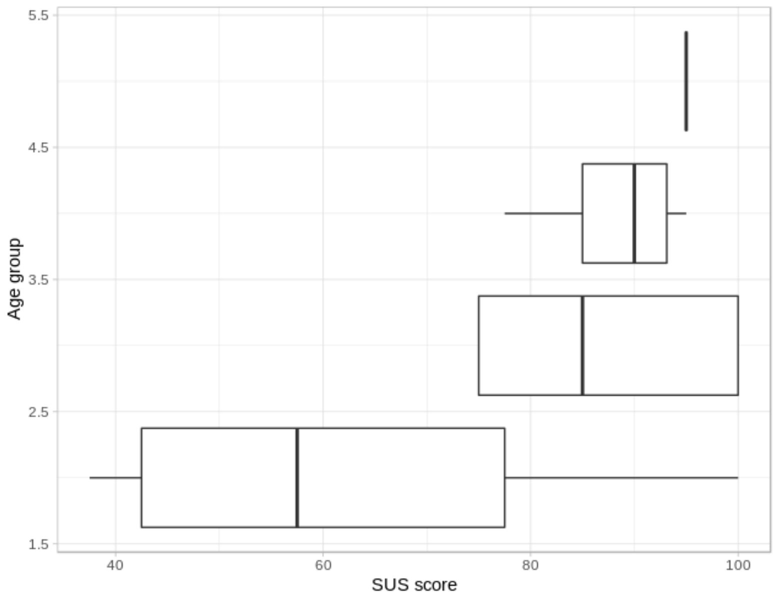
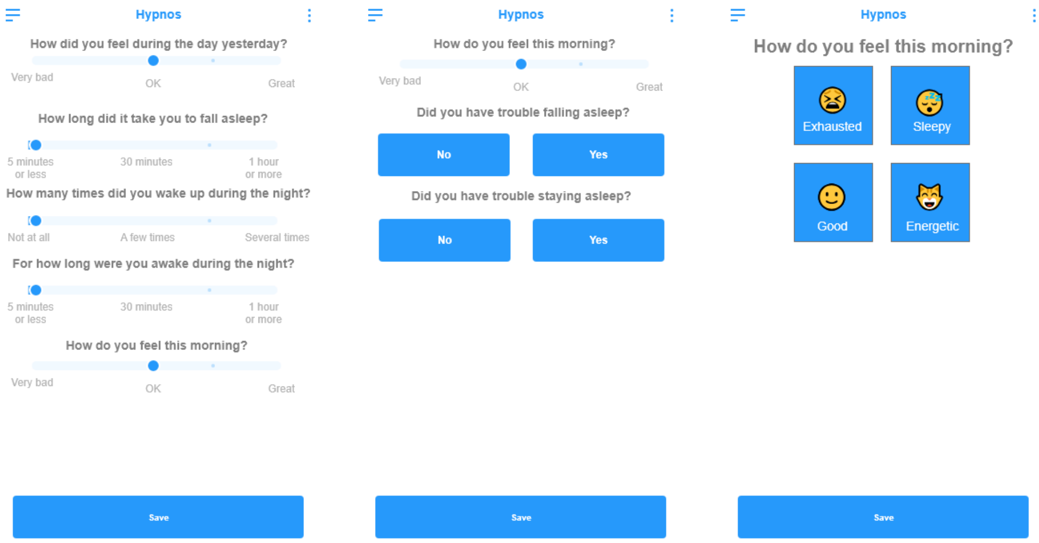


Figure 4.14.: Boxplot showing the distribution of SUS scores for the "fields" interface grouped by age

As presented in the correlation matrix in table 4.8, there are few strong correlations towards SUS scores. The biggest number is the moderate correlation between age and SUS score of "fields" ($r = 0.56$). As can be seen in figure 4.14, respondents in age group 2 (25-34) and upward were more positive to this interface. Otherwise, there are weak correlations between gender and the SUS score of "clock" ($r = 0.32$) and "fields" ($r = -0.29$) as well as a weak negative correlation between age and sleep tracking ($r = -0.35$). The is consistent with previous findings, where younger respondents have been more likely to have engaged in sleep tracking.



(a) (b) (c)

Figure 4.15.: Three wireframes that allows input of subjective sleep quality

4.3.2. Interface 2: Input of subjective sleep quality

Usability score

As can be seen in table 4.9 both the "Buttons" interface and the "smileys" interface had a mean 86.6 SUS score, while the "sliders" interface had a score of 76.5.

Despite having the same usability score, people responded with preferring the "buttons" interface when asked to select a interface. 13 responses preferred the "buttons" interface, and 6 preferred each of the other interfaces. This discrepancy may help us understand which usability factors are most important in this type of interface. For instance, "buttons" scored higher on the questions about frequency of use, the confidence of use, and integration of functions. "Smileys" scored better on ease of use, but worse on learnability.

Figure 4.16 shows a correlation matrix that includes the Pearson correlation between the different factors that were gathered. Noteworthy is the moderate correlations between the SUS score of the "sliders" interface and "smileys" interface (r = 0.55), and the SUS score of "sliders" and "buttons" (r = 0.41). Between the SUS scores of "sliders" and "buttons", there is only a weak correlation (r = 0.34), suggesting that more respondents rated one or the other higher, and not both. The respondents preferred interface is only weakly correlated with their SUS scores (at most r = -0.35). Use of sleep tracking is weakly correlated with a higher SUS score of the "Buttons" interface (r = 0.35), suggesting that there may be some preference of this interface by this user group. Gender negatively

Table 4.9.: Mean responses to SUS questionnaire and SUS score

Set 2	"Buttons"	"Sliders"	"Smileys"
I think that I would like to use this interface frequently	3.88	3.24	3.64
I found the interface unnecessarily complex	1.16	2.08	1.08
I thought the interface was easy to use	4.44	3.88	4.56
I think that I would need the support of a technical person to be able to use this interface	1	1.04	1
I found the various functions in this interface were well integrated	3.96	3.44	3.8
I thought there was too much inconsistency in this interface	1.68	1.44	1.6
I would imagine that most people would learn to use this interface very quickly	4.4	4.16	4.6
I found the interface very cumbersome to use	1.44	2.08	1.4
I felt very confident using the interface	4.32	3.88	4.28
I needed to learn a lot of things before I could get going with this interface	1.08	1.36	1.16
SUS	86.6	76.5	86.6
Percentage of participants preferred	52	24	24

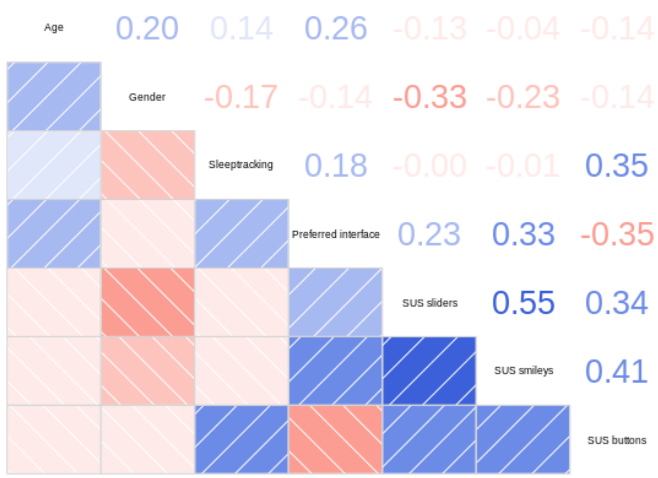
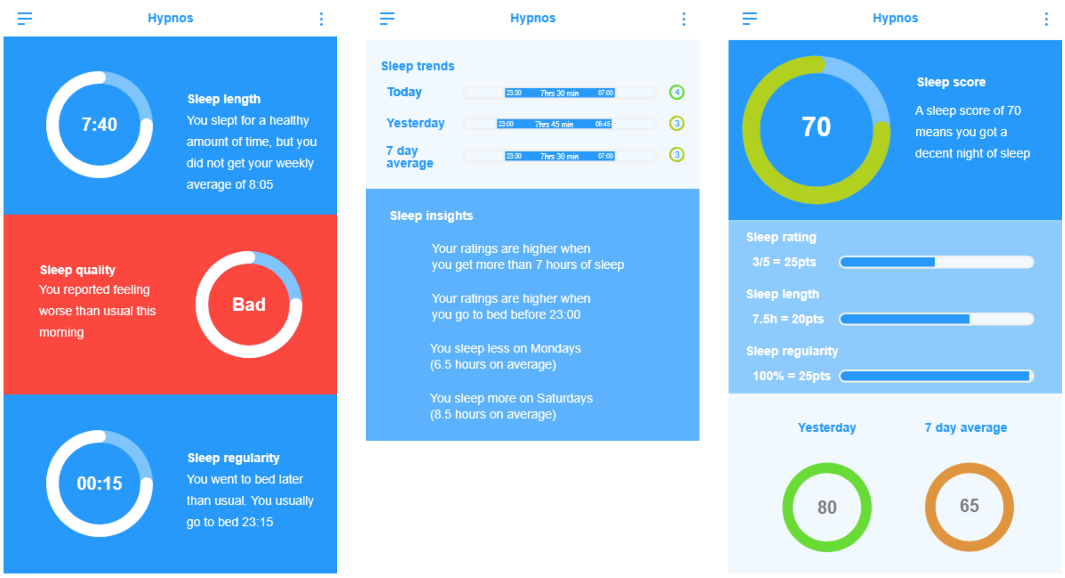


Figure 4.16.: Matrix showing the correlations between different factors and the SUS scores

correlates with the SUS score of "sliders" (r = -0.33).

4.3.3. Interface 3: Giving the user feedback on their sleep habits



(a) (b) (c)

Figure 4.17.: Three wireframes that gives the user feedback on their sleep habits

Table 4.10.: Mean responses to SUS questionnaire and SUS score

Set 3	"daily"	"insights"	"score"
I think that I would like to use this interface frequently	3.81	3.19	3.35
I found the interface unnecessarily complex	2.12	2.96	2.65
I thought the interface was easy to use	4	3.38	4.04
I think that I would need the support of a technical person to be able to use this interface	2.04	2.42	2.35
I found the various functions in this interface were well integrated	3.96	3.23	3.65
I thought there was too much inconsistency in this interface	2.31	2.62	2.54
I would imagine that most people would learn to use this interface very quickly	4.12	3.23	3.5
I found the interface very cumbersome to use	2.42	2.96	2.62
I felt very confident using the interface	3.96	3.46	3.65
I needed to learn a lot of things before I could get going with this interface	2.35	2.65	2.58
SUS	71.5	57.2	63.6
Percentage of participants preferred	42	4	54

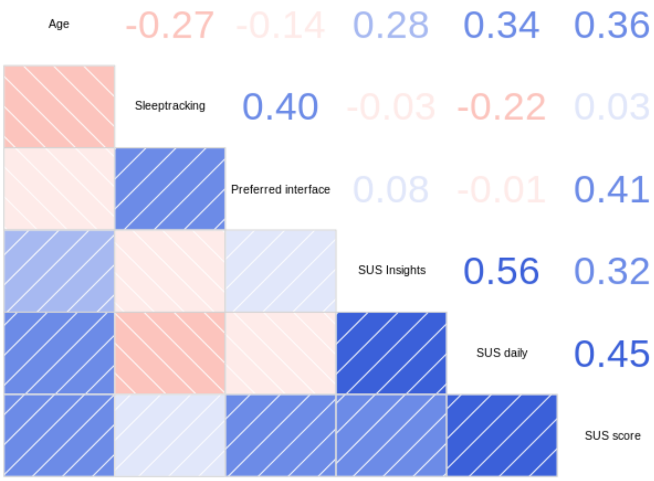


Figure 4.18.: Matrix showing the correlations between different factors and the SUS scores

Usability score

Here there is also a difference in which interface scored better on the SUS and which was preferred when asked directly. The "daily" interface scored the best with a SUS of 71.5. The most preferred was the "score" interface, with 54% of the votes and a SUS of 63.6. The "score" interface also scored worse on all but one of the individual questions, which was ease of use. This discrepancy between score and preference could indicate that there is a difference between these two interfaces that the questionnaire doesn't measure. One such difference could be aesthetics, which is not measured in the SUS questionnaire.

In the correlation matrix in figure 4.18, there is a moderate correlation between both the SUS of "insights" and "daily" ($r = 0.56$), and the SUS of "daily" and "score" ($r= 0.45$). Between "insights" and "score" it was lower ($r = 0.32$). Use of sleep tracking has a moderate correlation with the preferred interface ($r = 0.40$), so system experience is a factor of which interface was preferred. Interestingly, the preferred interface has a moderate correlation with the SUS of the "score" interface ($r = 0.41$), while it only has a very weak correlation with the other scores.

4.3.4. Summary

In the first usability test, seemingly the most complicated interface was the most preferred. This may be because it is more intuitive, as it is a time input interface with an interactive element resembling a clock. It may also be that it is more eye-catching, as it both uses more of the page and has more graphical elements. Younger respondents were more favourable to the more common smartphone interfaces, perhaps because they are more familiar with the technology.

In the second set of interfaces, which was for the input of subjective sleep quality, the most preferred interface was in the middle in terms of complexity. The results indicate that the most complex interface had too many interactive elements. Why the most straightforward interface was not preferred is unknown, but it may have been that it used aesthetics that may not increase perceived credibility (as shown by Oyibo, Adaji, and Vassileva [40]).

In the results for the set of interfaces meant to give the user feedback on their sleep, there was a discrepancy between the preferred interface selected by the participant and the one that got the better SUS score. This may be due to the preferred interface having some weaker aspects, but that it included a concept that the respondents preferred. In this case, this may be the sleep score that was included in the "score" interface.

5. Prototype

This chapter presents a high-fidelity prototype for a sleep tracking system, named *Hypnos*. This prototype is both a culmination of the findings from the user studies, but it is also an artefact that can be user-tested and further iterated upon. The act of creating a high-fidelity prototype can also be an excellent way to identify implementation difficulties and new opportunities.

The prototype was made in a somewhat iterative process. First, the *backend*, or the data structure, was developed before the wireframe evaluations were done. As the results from each wireframe evaluation came in, work started on the implementation while the next interface was evaluated. This proved to be a productive workflow. The prototype was informally tested with fellow students during development.

The prototype was built as a *progressive web application*, which is a web-based application with offline capabilities. They are essentially websites that can be added to the system level of many operating systems. Development was done using the React framework on top of the JavaScript programming language. Another framework that was used is *Bootstrap*, which is a design framework that makes it easier to create modern-looking applications. Where possible, pre-built solutions were used to reduce development time. This includes packages such as *LocalForage* and *Moment* that handle data storage and datetime formatting, as well as components such as *react-circular-input* that creates a functional circular input. These are all available through the *npm* package manager under open source licenses.

5.1. Modules

The prototype has two modules, one that allows the user to log their sleep and one that lets them view different representations of their data. In this chapter, the first module will be referred to as the "logging interface", and the second module as the "history interface".

Logging interface

The first module is the *logging* interface (figure 5.1), where the user can log their time spent sleeping, their perceived sleep quality, as well as if they had trouble falling asleep or staying asleep. This is data that can be used to identify sleep issues using established

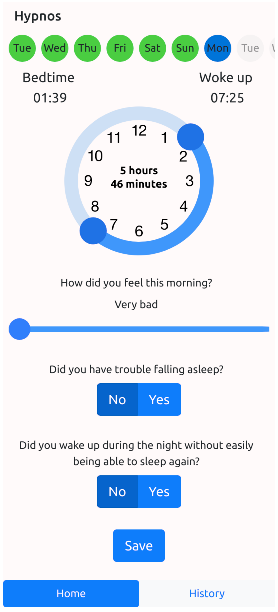


Figure 5.1.: Logging interface

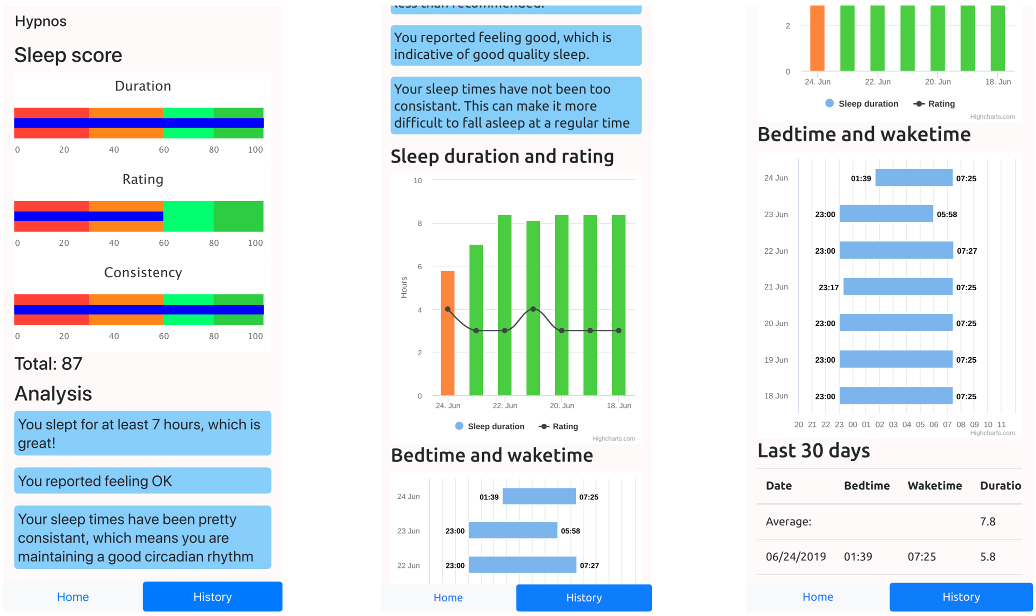
methods from sleep medicine. The manual nature of tracking this data can also act as a prompt for the user to reflect on their sleep.

The user can select any day from the previous week from a series of icons at the top of the screen. This icon will show as blue if it is selected and green if it has been logged already. Days from the following week are displayed with some translucency, with the intent to give the impression of a calendar. When selecting a previous day that has been logged, the inputs are updated to the saved data, allowing the user to view and edit the saved data. Currently, the standard bedtime and wake-up time is set somewhat arbitrarily, but in a final product, it could reflect a user-defined goal for sleep time or be based on recent trends.

The style of the inputs was selected based on the wireframe evaluations (see chapter 3 and 4). The bedtime and wake-up time selector was made by modifying an available *react component* to accept a second handle and display the range between them. The rest of the inputs were made using HTML form elements, using the CSS framework *Bootstrap* for a modern and consistent style. The save button is animated to indicate to the user that their data has been saved. This was done after several testers remarked that they were unsure if it had saved after pushing the button.

History interface

This module allows the user to view their historical data in several ways. Based on the results from the wireframe evaluations, a score-based system was chosen as the primary type of feedback. This implementation of a sleep score consists of three components: sleep duration, sleep quality and consistency. These components are each given a score from 0 to 100, which is averaged into the total score. Each component is displayed as a bar graph where the background goes from red to orange, and then to green. The colour attempts to associate the score with a degree of desirability.



(a) (b) (c)

Figure 5.2.: The history interface

Since interviews showed that users of sleep tracking systems wanted personalised feedback, there was an attempt to create text-based feedback to the previous night's sleep. The current implementation consists of feedback on the length of sleep, the user's rating of their sleep and their consistency of sleep. It is not particularly personalised, as the feedback is based on general recommendations. In a final product, one could imagine that the system would start with these general recommendations, before establishing a baseline for each user. While this is information that is available in the sleep score, the different presentation is worth exploring, as it may appeal more to certain users. It could potentially present more information, that can further explain the reasoning behind the score or propose interventions if a negative trend is detected. Detecting anomalies, such as one night of poor sleep, could be used to provide reassurance or coping strategies.

Experiments were also made using charts to visualise different aspects of sleep. First, there is one chart that shows both the time spent sleeping as well as the rating. This could potentially help the user find a relationship between the amount of sleep and how rested they feel when they wake up. Current systems, such as Fitbit, do not have this direct comparison; the user has to look at two different charts in order to compare data. Next is a chart showing the duration of each night of sleep. This type of chart could be useful for visualising consistency of sleep, and to show how circadian rhythms can affect sleep.

5.2. Future development

This section outlines possible features that could be part of future iterations of *Hypnos*.

Supporting maintenance As established by previous research [43], there is a lack of support for maintaining user behaviour in personal tracking systems. One way of supporting this could be for the system to detect changes over time or identify periods where sleep is particularly good or bad, as suggested by Choe et al. [10]. The system already tracks subjective sleep quality and indicators of insomnia, which can be used establish a baseline and to identify nights that diverts from it. Such functionality could assist the user in keeping up their habits and identify factors that could affect them, such as diet, stress and activity. If these factors also are being tracked, the system could attempt to automatically identify if changes in one factor lead to changes in sleep quality or quantity. This sort of personalised feedback has been a frequent request, both from researchers [43, 28] and from participants in the user studies presented in chapter 4.

Ravichandran et al. [43] recommends that sleep tracking systems should include subjective sleep quality assessment, focus on actionable feedback, and give feedback in ranges. This prototype includes a subjective sleep quality assessment, and it gives feedback in ranges via the sleep score implementation. It may not quite qualify as giving actionable feedback, although the text-based feedback may develop in that direction.

Exploring different types of reflection As seen in chapter 2, some systems provide learning modules, namely the system SHUTi, in which this is the focus. One direction to take this prototype is to include a learning module, that ensures that the user has an understanding of what healthy sleep is, and what they can do in order to improve their sleep. Liang and Ploderer [28] showed that many users were unsure if their sleep was healthy and that they did not know how to act. A well-designed learning module could help alleviate these barriers. This learning module should include both information on healthy sleep, but also tools that help the user deal with issues such as insomnia and social jetlag.

Integration with other services Many of the current fitness trackers have methods for sharing data via an API, which makes it possible to collect data from these sources. When it comes to sleep data, one question is to which degree one should make it easier for the user to input data. If one were to gather data from an external source automatically, the application could lose some of its reflective nature as the user is no longer required to interact with it. A middle ground could be to use these data to populate the journal but still require the user to look over it and possibly modify it. Keeping subjective evaluations would be critical to an experience that leans towards CBT-I, and may also be more reliable than some current sleep tracking systems [1]. On the other hand, users in our survey have ranked automatic tracking and sleep stage tracking as essential features of sleep tracking.

The possibility to gather data from other areas of health such as diet and exercise is alluring, but some considerations must be taken. Depending on how it was tracked, it may not be entirely representative of the real world. Someone who tracks their running may not track other activities such as walking or swimming. Step counts do not track cycling or weight lifting. With diet tracking, there are people who do not track every day, or perhaps not every meal. It can, therefore, be problematic to attempt to correlate these data with sleep data if the goal is to make meaningful feedback to the user.

Alternative mediums This prototype was made to work on smartphone devices primarily. But it is possible to imagine sleep tracking systems that use different interfaces. Many wearable sleep tracking systems have both a smartphone view and a view on the device itself. It would be a different task to create a sleep tracking system for a smartwatch, as the interaction model would have to change.

Voice control is another possible interaction model that has not been explored within sleep tracking. It is possible to add to existing systems, which could help with accessibility for some users, but it could also open up new user experiences. Instead of using their smartphones or smartwatches to log their sleep, users could talk to *smart speakers* such as Google Home. One could even imagine a *smart home* sleep tracker that uses motion detecting sensors in order to track sleep. By connecting with *smart home* devices it could keep track of temperatures, meals eaten and even time spent watching TV.

6. Conclusions and future work

This chapter will conclude the thesis by answering the research questions that were established in the introduction. Further, it will identify future research and development opportunities.

How users interact with sleep tracking systems (RQ1) In the survey, there were both users of wearable sleep tracking systems and smartphone systems. While they were almost even in terms of the number of users (18 vs 17), the demographic was quite different. Smartphone systems were primarily used by the youngest demographic (18-24), while users of wearable systems tended to be older. This is an indication that smartphone sleep tracking systems may appeal more to a younger demographic.

The survey also showed that many interact with their sleep data on close to a daily basis. There is also a substantial group that do this weekly. The interviews showed similar tendencies and gave the impression that those who check less often are more interested in evaluating trends. Sleep tracking systems should attempt to cater to both sets of users by providing interfaces that fit their needs.

The interviews suggest that users of sleep tracking systems often interact with their data in the morning. Designers of sleep tracking systems should keep this in mind when considering to add features that require the user to interact at different times of day.

User motivations for using sleep tracking systems (RQ2) From the interviews, three groups of users stood out. Many were *goal driven* and wanted to improve their sleep or fitness. Some focused on reaching a total amount of sleep, while others had goals related to sleep quality. Other users were *health driven*. These users used their sleep tracking system as a measurement of their health. One interviewed user had changed medications based on their sleep data. *Technology driven* users were motivated by the technology itself and did not have any set goals. Additionally, the survey showed one substantial user group: those that gave "smart alarm" as a reason for using a sleep tracker. 12 of 36 sleep tracking users gave this as a reason, which is considerable. As this is mainly a feature of smartphone systems it may be wise for developers of wearable systems to include this feature.

Necessary functions of sleep tracking systems (RQ3) The survey asked participants to rate the importance of different features of sleep tracking systems. Features such as "automatic tracking" and "show time spent in sleep stages" were highly rated. "Reminder to go to bed" was rated the lowest. This was also the case when the sample was divided by sleep tracking use, but this also showed that sleep tracking users rated certain functions higher. "Alarm function" and "ability to enter additional data" did better with the sleep trackers, but "sleep score" and "advice" did worse. It could be that users of sleep tracking have seen these functions already and did not enjoy them.

From the interviews, it became clear that users used both the daily view and longer-term views. They used a daily view to study sleep quality and awakenings. The longer-term view they would use to make comparisons, either by periods of time or with other data such as exercise and diet. Supporting such comparisons should be considered a necessary feature. The prototype currently supports comparing time spent sleeping with a subjective rating. The survey showed that exercise and diet tracking was common among sleep trackers, and as they both can influence sleep they may be integrated into such a comparison.

Some current systems do give some advice, specifically Fitbit with their "sleep insights". Fitbit users that were interviewed did not feel that they were that personalised. Users were unsure of why they got the advice they got, and they felt that it was too generic. One user was not sure why the system told him to get a new pillow and thought it was product placement. Nevertheless, they all agreed that a more suitable version of this would be a useful feature. The prototype shows how such advice could be displayed, although the proper algorithms have to be developed in order to generate interesting advice.

The wireframe evaluations showed that the participants favoured more straightforward input methods, but did not want them to be too simple. This suggests that they too see the value of capturing more data. In the "feedback" interface, user preference was towards the interface that included a sleep score. This could mean that a sleep score is preferred to text-based advice, which was one component of the other wireframes. The current prototype includes both, and could be modified in order to test this hypothesis.

How sleep tracking systems can better support healthy sleep behaviour (RQ4) The medical research shows that consistency, both in the amount of sleep and timing of sleep, is an essential factor of healthy sleep. Current systems may reward quantity and quality of sleep, but rewarding the consistency of sleep timing has not been done. In the prototype developed, this was done by including consistency as a factor of a total sleep score.

Interviews showed that users want personalised advice. It also showed that such advice should include explanations for why it is given, and why it can help the user in the given situation. It should not be too generic and also be something that the user can realistically do.

Sleep tracking systems should attempt to provide learning experiences, perhaps via advice or by including a learning module. Concepts such as *chronotypes* and *social jetlag* can be used. Previous research has shown that many users have different views on sleep than does medical science, which was also an impression given by the interviews. For the technology of sleep tracking to reach its potential of helping users sleep better, more user education is needed.

6.1. Future work

In order to assess the performance of the prototype, it should be user tested. This can be done both quantitatively and qualitatively. As with the wireframe evaluations, it could be compared to other sleep tracking systems using the SUS questionnaire, which may be better suited to functional prototypes. Task-oriented experiments are another possibility. Qualitative research such as observational studies, diary studies and case studies can also be used to gain user opinions regarding the functionality of the prototype. These methods can be combined with additional interviews and surveys in order to get more information on the user experience.

As outlined in section 5.2, several possible features could be added to the prototype in the future. This development could be structured using an iterative process, where a feature is added, before being user-tested to evaluate what it adds to the user experience. For *Hypnos* to become a real product, it must also establish a business model in order to become self-sustainable. In future research, this could be a topic of investigation, which could include how many users are willing to pay for the product, whether a subscription-based model is suitable, and how to market the product.

Appendices

A. Interview information

Request to participate in a study of digital sleep aids

This is a request for you to participate in a research project where the purpose is to develop a prototype of a new application that will help reflect on sleep patterns. In this letter we give you information about the goals of the project and what participation will mean for you.

Purpose
For a master's thesis in media and interaction design, the goal is to develop a prototype that will be able to help reflect on sleep patterns and give advice that can provide increased sleep quality. As part of this project, we want to get in touch with users of existing products in order to be able to map the motivation behind the use and how they are actually used. The result of these surveys will inform the design of the prototype and be included in the master thesis.

Who is responsible for the research project?
The project is part of a master's thesis under the auspices of the Department of Information and Media Studies at the University of Bergen.

Why do you have questions about participating?
You are asked to participate because you have shown interest in participating through social media or other postings.

What does it mean for you to participate?
If you choose to participate in the project, this means that you participate in an interview. It will take you approx. 30 minutes. The interview contains questions about your use of digital aids whose purpose is to register or in other ways reflect on sleep. None of the questions will directly address your health, but if you choose to enter the subject yourself, I ask for consent to process the information. The interview may be recorded with audio recorder, and later transcribed by interviews. The content itself will not be linked to your contact information or person.

Excerpts from the interview may be published in the Master's thesis. These will be anonymised and rewritten to avoid sensitive information.

You may be asked to participate in a follow-up study. Your contact information will then be registered in a separate list.

Participation
It is voluntary to participate in the project. If you choose to participate, you may at any time withdraw your consent without giving any reason. All information about you will then be anonymized. It will have no negative consequences for you if you do not want to participate or later choose to withdraw.

Your privacy - how we store and use your information
We will only use the information about you for the purposes we have told about in this written. We treat the information confidentially and in accordance with the privacy policy.

- The information will only be processed by the student.
- Names and contact information will be kept separate from notes and audio recordings.

What happens to your information when we close the research project?
The project is scheduled to end on 30.06.2019. Contact information, notes and sound recordings will then be deleted. Only anonymised extracts relevant to the master thesis will be published.

Your rights
As long as you can be identified in the data material, you have the right to:
- know what personal information is registered about you
- to have your personal information corrected
- to have your personal information deleted
- a copy of your personal data (data portability)
- to send a complaint to the Data Protection Officer or the Data Inspectorate about the processing of your personal data.

What gives us the right to process personal information about you?
We process information about you based on your consent.

On behalf of the Department of Information and Media Studies NSD - The Norwegian Center for Research Data AS has considered that the processing of personal data in this project is in accordance with the privacy policy.

Where can I find out more?
If you have questions about the study *Development of sleep system*, or want to use your rights, please contact:
- Master student - Gøran Slettemark by email (goran.slettemark@uib.no) or phone: 97 48 97 50)
- Supervisor - Christoph Trattner by email (christoph.trattner@uib.no) or phone: +43 65 0 2 40 28 01
- NSD - Norwegian Center for Research Data AS, by e-mail (personvernombudet@nsd.no) or by phone: 55 58 21 17.

Sincerely

Gøran A. Slettemark

Project Manager

--- -- -------------------------

Declaration of consent

I have received and understood information about the projectand have been given the opportunity to ask questions. I agree to:

- □ participate in the interview
- □ have the interview recorded

I agree that my information is processed until the project is completed, approx. 30.06.2019

--- -- ---------------
(Signed by project participant, date)

B. Assessment from NSD

NSD Personvern

16.11.2018 11:14

Det innsendte meldeskjemaet med referansekode 252690 er nå vurdert av NSD.

Følgende vurdering er gitt:

Det er vår vurdering at behandlingen vil være i samsvar med personvernlovgivningen, så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet 16.11.2018 med vedlegg, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD ENDRINGER

Dersom behandlingen av personopplysninger endrer seg, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. På våre nettsider informerer vi om hvilke endringer som må meldes. Vent på svar før endringen gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET

Prosjektet vil behandle særlige kategorier av personopplysninger om helse og alminnelige personopplysninger frem til 30.06.2019.

LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 nr. 11 og art. 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse, som kan dokumenteres, og som den registrerte kan trekke tilbake.

Lovlig grunnlag for behandlingen vil dermed være den registrertes uttrykkelige samtykke, jf. personvernforordningen art. 6 nr. 1 a), jf. art. 9 nr. 2 bokstav a, jf. personopplysningsloven § 10, jf. § 9 (2).

PERSONVERNPRINSIPPER

NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen:

- om lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til nye uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).
NSD vurderer at informasjonen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER

NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

For å forsikre dere om at kravene oppfylles, må prosjektansvarlig følge interne retningslinjer/rådføre seg med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET

NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Kajsa Amundsen
Tlf. Personverntjenester: 55 58 21 17 (tast 1)

C. Interview guide

Information

- What the interview is used for
- Anonymity and duty of confidentiality
- Inform about audio recordings
- Get permission

Background

- Age
- Profession

Main part

- What kind of sleep tracker do you use?
  - How does it work?
  - Have you used it for a long time?
  - Why did you get it?
- Are you using it every day?
- Do you often look at the data?
  - When during the day?
  - How often?
- What are you looking for?
  - Fitbit: Sleep Insights
- Does the measurement affect your day? Eg. training or bedtime?
- Do you look at trends?
  - Link them to other things?
  - Have you changed anything based on your data?
- Have you used something similar earlier or for other things?
  - What did they do different?
  - Why don't you use it anymore?
- Do you track other things?
- What do you get from using this?
- What is missing?
- What would you change?

Summary

- Summarise findings
- "Anything you want to add?"

D. Survey questionnaire

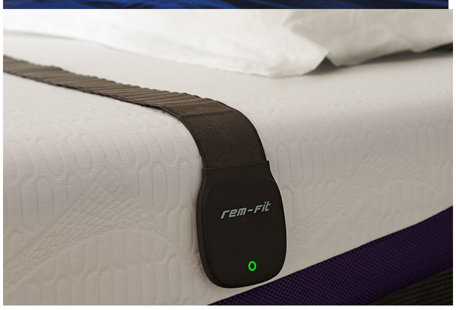
What is your age?
- ☐ 18-24 years
- ☐ 25-34 years
- ☐ 35-44 years
- ☐ 45-54 years
- ☐ 55-64 years
- ☐ 65 years or older

What is your gender?
- ☐ Male
- ☐ Female
- ☐ Other_________

Do you use a sleep tracker application or device?
- ☐ Yes
- ☐ No

What kind of sleep tracker do you use?
- ☐ Smartwatch (Apple Watch, FitBit, Garmin, etc.)
- ☐ Smartphone app (Sleep Cycle, Sleep as Android, SnoreLab, etc.)
- ☐ Dedicated sleep tracking device (Withings Aura, Beddit, etc.)
- ☐ Other (specify): _________

Examples of dedicated sleep tracking devices:



On the left is a dedicated bedside sleep tracker. To the right is an in-bed dedicated sleep tracker

For how long have you used this sleep tracker?

- ☐ Less than a month
- ☐ Less than 3 months
- ☐ Less than a year
- ☐ More than a year

How many days a week do you look at your sleep data?

- ☐ 0
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7

What are the main reasons you use a sleep tracker?

- ☐ To understand my sleeping behaviour
- ☐ Alarm function (smart alarm, gentle wake-up)
- ☐ Remind me to go to bed
- ☐ Other (specify)_________

Do you otherwise keep track of ... ?

- ☐ Diet
- ☐ Exercise or activity
- ☐ Mood
- ☐ Finance
- ☐ Habits (Stopping smoking, productivity)
- ☐ Health related measurements (Heart rate, blood sugar, etc.)
- ☐ Other: _________

How familiar are you with the following?

	1 - Not heard of	2	3	4	5 - Very familiar
Activity tracking (Strava, Runkeeper, JeFit, etc)	☐	☐	☐	☐	☐
Diet tracking (MyFitnessPal, FatSecret, Cronometer, etc)	☐	☐	☐	☐	☐
Habit tracking (Habitica, Streaks, Momentum, etc)	☐	☐	☐	☐	☐

Do you wear a smartwatch?

- ☐ Yes
- ☐ No

Why do you wear a smartwatch?

- ☐ To get notifications
- ☐ Measure activity
- ☐ Sleep tracking
- ☐ To use smartphone applications on my watch
- ☐ Other: ________

How do you agree with the following statements?

	1 - Strongly disagree	2	3	4	5 - Strongly agree
I plan to use a sleep tracker in the future	☐	☐	☐	☐	☐
A sleep tracker would help me sleep better	☐	☐	☐	☐	☐
A sleep tracker should estimate my quality of sleep	☐	☐	☐	☐	☐
A sleep tracker should show me how long i have slept	☐	☐	☐	☐	☐
A sleep tracker should give me a score based on my sleep quality and length	☐	☐	☐	☐	☐
A sleep tracker should tell me when I should go to bed	☐	☐	☐	☐	☐

How do you agree with the following statements?

	1 - Strongly disagree	2	3	4	5 - Strongly agree
I want to see how exercise impacts my sleep	☐	☐	☐	☐	☐
I want to see how activity (walking, standing, etc) impacts my sleep	☐	☐	☐	☐	☐
I want to see if my diet impacts my sleep	☐	☐	☐	☐	☐
I want to know if my phone use impacts my sleep	☐	☐	☐	☐	☐
I want to know if noise impacts my sleep	☐	☐	☐	☐	☐
I want to know if stress impacts my sleep	☐	☐	☐	☐	☐
I want to know if I sleep better when I spend more time outside	☐	☐	☐	☐	☐
I want advice on how to sleep better	☐	☐	☐	☐	☐

How important are these features in a sleep tracker?

	1 - Not important	2	3	4	5 - Very important
Advice on sleep habits	☐	☐	☐	☐	☐
Integration with diet and exercise trackers	☐	☐	☐	☐	☐
Ability to enter additional data	☐	☐	☐	☐	☐
Alarm function	☐	☐	☐	☐	☐
Automatic tracking	☐	☐	☐	☐	☐
A daily 'sleep score'	☐	☐	☐	☐	☐
Ability to rate how tired you feel	☐	☐	☐	☐	☐
Reminder to go to bed	☐	☐	☐	☐	☐
Show time spent in sleep stages	☐	☐	☐	☐	☐
Showing trends	☐	☐	☐	☐	☐

Are there other features that are important to you?

____

E. Prototype information

The current prototype is available at *https://hypnosa.herokuapp.com/*, and the source code is available on GitHub *https://github.com/goransle/hypnos*.

A few notes on the prototype:

- It is best viewed on a smartphone, or by using the mobile view in the developer tools in your favourite browser.
- It saves to the local storage of the browser, so saved data will not transfer from device to device, or to a different browser.
- As the prototype is hosted on a free service, it may require some spin-up time if no one has used in a while.
- Using Google Chrome on Android, you can add Hypnos to the home screen by selecting *"Add to Home screen"* in the menu. In Firefox, this option is found under *Page*, and is called *"Add Page Shortcut"*.

Bibliography

[1] Torbjörn Åkerstedt et al. "The Meaning of Good Sleep: A Longitudinal Study of Polysomnography and Subjective Sleep Quality". In: *Journal of Sleep Research* 3.3 (1994), pp. 152–158. ISSN: 1365-2869. DOI: 10.1111/j.1365-2869.1994.tb00122.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2869.1994.tb00122.x (visited on 01/28/2019).

[2] Siobhan Banks. "Behavioral and Physiological Consequences of Sleep Restriction". In: *Journal of Clinical Sleep Medicine* 3.5 (2007), p. 10.

[3] Kelly Glazer Baron and Kathryn J. Reid. "Circadian Misalignment and Health". In: *International Review of Psychiatry* 26.2 (Apr. 1, 2014), pp. 139–154. ISSN: 0954-0261. DOI: 10.3109/09540261.2014.911149. pmid: 24892891. URL: https://doi.org/10.3109/09540261.2014.911149 (visited on 04/08/2019).

[4] Jared S. Bauer et al. "ShutEye: Encouraging Awareness of Healthy Sleep Recommendations with a Mobile, Peripheral Display". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1401–1410.

[5] Bjørn Bjorvatn. *Søvndagbok SOVno*. URL: https://helse-bergen.no/nasjonal-kompetansetjeneste-for-sovnsykdommer-sovno/sovndagbok-sovno (visited on 05/22/2019).

[6] Jan Born, Björn Rasch, and Steffen Gais. "Sleep to Remember". In: *The Neuroscientist* 12.5 (Oct. 1, 2006), pp. 410–424. ISSN: 1073-8584. DOI: 10.1177/1073858406292647. URL: https://doi.org/10.1177/1073858406292647 (visited on 05/20/2019).

[7] John Brooke. "SUS: A Retrospective". In: *Journal of usability studies* 8.2 (2013), pp. 29–40.

[8] Daniel J. Buysse et al. "The Pittsburgh Sleep Quality Index: A New Instrument for Psychiatric Practice and Research". In: *Psychiatry Research* 28.2 (May 1, 1989), pp. 193–213. ISSN: 0165-1781. DOI: 10.1016/0165-1781(89)90047-4. URL: http://www.sciencedirect.com/science/article/pii/0165178189900474 (visited on 04/08/2019).

[9] Christian Cajochen et al. "Evening Exposure to a Light-Emitting Diodes (LED)-Backlit Computer Screen Affects Circadian Physiology and Cognitive Performance". In: *Journal of Applied Physiology* 110.5 (Mar. 17, 2011), pp. 1432–1438. ISSN: 8750-7587. DOI: 10.1152/japplphysiol.00165.2011. URL: https://www.physiology.org/doi/full/10.1152/japplphysiol.00165.2011 (visited on 04/08/2019).

[10] Eun Kyoung Choe et al. "Opportunities for Computing Technologies to Support Healthy Sleep Behaviors". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* (Vancouver, BC, Canada). CHI '11. New York, NY, USA: ACM, 2011, pp. 3053–3062. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1979395. URL: http://doi.acm.org/10.1145/1978942.1979395 (visited on 04/08/2019).

[11] Eun Kyoung Choe et al. "SleepTight: Low-Burden, Self-Monitoring Technology for Capturing and Reflecting on Sleep Behaviors". In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* (Osaka, Japan). UbiComp '15. New York, NY, USA: ACM, 2015, pp. 121–132. ISBN: 978-1-4503-3574-4. DOI: 10.1145/2750858.2804266. URL: http://doi.acm.org/10.1145/2750858.2804266 (visited on 04/08/2019).

[12] Jason Cipriani. *Fitbit's Sleep Score Beta Looks Promising.* Dec. 1, 2019. URL: https://www.zdnet.com/article/fitbits-sleep-score-beta/ (visited on 06/03/2019).

[13] Nediyana Daskalova et al. "SleepCoacher: A Personalized Automated Self-Experimentation System for Sleep Recommendations". In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology.* (Tokyo, Japan). UIST '16. New York, NY, USA: ACM, 2016, pp. 347–358. ISBN: 978-1-4503-4189-9. DOI: 10.1145/2984511.2984534. URL: http://doi.acm.org/10.1145/2984511.2984534 (visited on 04/08/2019).

[14] Fitbit. *How Do I Track My Sleep with My Fitbit Device?* 2019. URL: https://help.fitbit.com/articles/en_US/Help_article/1314 (visited on 06/03/2019).

[15] Fitbit. *REM, Light, Deep: How Much of Each Stage of Sleep Are You Getting?* Sept. 19, 2018. URL: https://blog.fitbit.com/sleep-stages-explained/ (visited on 06/03/2019).

[16] Brian J. Fogg. "Persuasive Computers: Perspectives and Research Directions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM Press/Addison-Wesley Publishing Co., 1998, pp. 225–232.

[17] Thomas Fritz et al. "Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2014, pp. 487–496.

[18] Janne Grønli and Reidun Ursin. *Basale søvnmekanismer.* URL: https://tidsskriftet.no/2009/09/tema-sovn/basale-sovnmekanismer (visited on 05/29/2019).

[19] Susanne Hagatun et al. "The Short-Term Efficacy of an Unguided Internet-Based Cognitive-Behavioral Therapy for Insomnia: A Randomized Controlled Trial With a Six-Month Nonrandomized Follow-Up". In: *Behavioral Sleep Medicine* 17.2 (Mar. 4, 2019), pp. 137–155. ISSN: 1540-2002. DOI: 10.1080/15402002.2017.1301941. pmid: 28345961. URL: https://doi.org/10.1080/15402002.2017.1301941 (visited on 04/08/2019).

[20] Sajanee Halko and Julie A. Kientz. “Personality and Persuasive Technology: An Exploratory Study on Health-Promoting Mobile Applications”. In: *Persuasive Technology*. Ed. by Thomas Ploug, Per Hasle, and Harri Oinas-Kukkonen. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2010, pp. 150–161. ISBN: 978-3-642-13226-1.

[21] Max Hirshkowitz et al. “National Sleep Foundation’s Sleep Time Duration Recommendations: Methodology and Results Summary”. In: *Sleep Health: Journal of the National Sleep Foundation* 1.1 (Mar. 1, 2015), pp. 40–43. ISSN: 2352-7218. DOI: 10.1016/j.sleh.2014.12.010. pmid: 29073412. URL: https://www.sleephealthjournal.org/article/S2352-7218(15)00015-7/abstract (visited on 05/21/2019).

[22] *How Does Oura Ring Measure Your Sleep Score*. URL: https://ouraring.com/sleep-score/ (visited on 06/17/2019).

[23] Michael Irwin et al. “Partial Night Sleep Deprivation Reduces Natural Killer and Cellular Immune Responses in Humans.” In: *The FASEB journal* 10.5 (1996), pp. 643–653.

[24] John G. Jenkins and Karl M. Dallenbach. “Obliviscence during Sleep and Waking”. In: *The American Journal of Psychology* 35.4 (1924), pp. 605–612. ISSN: 0002-9556. DOI: 10.2307/1414040. URL: https://www.jstor.org/stable/1414040 (visited on 05/22/2019).

[25] Murray W. Johns. “A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale”. In: *Sleep* 14.6 (Nov. 1, 1991), pp. 540–545. ISSN: 0161-8105. DOI: 10.1093/sleep/14.6.540. URL: https://academic.oup.com/sleep/article/14/6/540/2742871 (visited on 06/17/2019).

[26] Kristen L. Knutson et al. “The Metabolic Consequences of Sleep Deprivation”. In: *Sleep Medicine Reviews* 11.3 (June 1, 2007), pp. 163–178. ISSN: 1087-0792. DOI: 10.1016/j.smrv.2007.01.002. URL: http://www.sciencedirect.com/science/article/pii/S1087079207000202 (visited on 05/22/2019).

[27] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research Methods in Human-Computer Interaction*. Morgan Kaufmann, 2017.

[28] Zilu Liang and Bernd Ploderer. “Sleep Tracking in the Real World: A Qualitative Study into Barriers for Improving Sleep”. In: *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. ACM, 2016, pp. 537–541.

[29] Matthew D. Milewski et al. “Chronic Lack of Sleep Is Associated with Increased Sports Injuries in Adolescent Athletes”. In: *Journal of Pediatric Orthopaedics* 34.2 (2014), pp. 129–133.

[30] Charles M Morin, Sylive Rodrigue, and Hans Ivers. “Role of Stress, Arousal, and Coping Skills in Primary Insomnia | Ovid”. In: *Psychosomatic Medicine* 65.2 (2003), pp. 259–267. URL: https://oce.ovid.com/article/00006842-200303000-00012/HTML (visited on 05/29/2019).

[31] Charles M. Morin et al. "Psychological And Behavioral Treatment Of Insomnia: Update Of The Recent Evidence (1998–2004)". In: *Sleep* 29.11 (Nov. 1, 2006), pp. 1398–1414. ISSN: 0161-8105. DOI: 10.1093/sleep/29.11.1398. URL: https://academic.oup.com/sleep/article/29/11/1398/2709225 (visited on 06/07/2019).

[32] Hendrik Müller, Aaron Sedley, and Elizabeth Ferrall-Nunge. "Survey Research in HCI". In: *Ways of Knowing in HCI*. Ed. by Judith S. Olson and Wendy A. Kellogg. New York, NY: Springer New York, 2014, pp. 229–266. ISBN: 978-1-4939-0378-8. DOI: 10.1007/978-1-4939-0378-8_10. URL: https://doi.org/10.1007/978-1-4939-0378-8_10 (visited on 06/18/2019).

[33] Michael Muller. "Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method". In: *Ways of Knowing in HCI*. Ed. by Judith S. Olson and Wendy A. Kellogg. New York, NY: Springer New York, 2014, pp. 25–48. ISBN: 978-1-4939-0378-8. DOI: 10.1007/978-1-4939-0378-8_2. URL: https://doi.org/10.1007/978-1-4939-0378-8_2 (visited on 06/21/2019).

[34] S. A. Munson and S. Consolvo. "Exploring Goal-Setting, Rewards, Self-Monitoring, and Sharing to Motivate Physical Activity". In: *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. 2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops. May 2012, pp. 25–32. DOI: 10.4108/icst.pervasivehealth.2012.248691.

[35] Sean Munson. *Mindfulness, Reflection, and Persuasion in Personal Informatics*.

[36] Alain Muzet. "Environmental Noise, Sleep and Health". In: *Sleep Medicine Reviews* 11.2 (Apr. 1, 2007), pp. 135–142. ISSN: 1087-0792. DOI: 10.1016/j.smrv.2006.09.001. URL: http://www.sciencedirect.com/science/article/pii/S1087079206001055 (visited on 06/06/2019).

[37] *Notes on User Centered Design Process (UCD)*. URL: https://www.w3.org/WAI/EO/2003/ucd (visited on 03/29/2019).

[38] Lars Nyre. "Media Design Method". In: *The Journal of Media Innovations* 1.1 (2014), pp. 86–109.

[39] *Oura – Apps on Google Play*. URL: https://play.google.com/store/apps/details?id=com.ouraring.oura&hl=en_GB (visited on 06/10/2019).

[40] Kiemute Oyibo, Ifeoma Adaji, and Julita Vassileva. "What Drives the Perceived Credibility of Health Apps: Classical or Expressive Aesthetics". In: *HealthRecSys Workshop., Vancouver, Canada*. 2018.

[41] Dale E. Rae et al. "One Night of Partial Sleep Deprivation Impairs Recovery from a Single Exercise Training Session". In: *European Journal of Applied Physiology* 117.4 (Apr. 1, 2017), pp. 699–712. ISSN: 1439-6327. DOI: 10.1007/s00421-017-3565-5. URL: https://doi.org/10.1007/s00421-017-3565-5 (visited on 04/08/2019).

[42] Björn Rasch and Jan Born. “About Sleep’s Role in Memory”. In: *Physiological Reviews* 93.2 (Apr. 2013), pp. 681–766. ISSN: 0031-9333. DOI: 10.1152/physrev.00032.2012. pmid: 23589831. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3768102/ (visited on 06/17/2019).

[43] Ruth Ravichandran et al. “Making Sense of Sleep Sensors: How Sleep Sensing Technologies Support and Undermine Sleep Health”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* (Denver, Colorado, USA). CHI ’17. New York, NY, USA: ACM, 2017, pp. 6864–6875. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025557. URL: http://doi.acm.org/10.1145/3025453.3025557 (visited on 04/08/2019).

[44] Lee M. Ritterband et al. “Effect of a Web-Based Cognitive Behavior Therapy for Insomnia Intervention With 1-Year Follow-up: A Randomized Clinical Trial”. In: *JAMA Psychiatry* 74.1 (Jan. 1, 2017), pp. 68–75. ISSN: 2168-622X. DOI: 10.1001/jamapsychiatry.2016.3249. URL: https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2589161 (visited on 06/07/2019).

[45] Till Roenneberg et al. “A Marker for the End of Adolescence”. In: *Current Biology* 14.24 (Dec. 29, 2004), R1038–R1039. ISSN: 0960-9822. DOI: 10.1016/j.cub.2004.11.039. URL: http://www.sciencedirect.com/science/article/pii/S0960982204009285 (visited on 05/29/2019).

[46] Till Roenneberg et al. “Social Jetlag and Obesity”. In: *Current Biology* 22.10 (May 22, 2012), pp. 939–943. ISSN: 0960-9822. DOI: 10.1016/j.cub.2012.03.038. URL: http://www.sciencedirect.com/science/article/pii/S0960982212003259 (visited on 04/08/2019).

[47] John Rooksby et al. “Personal Tracking as Lived Informatics”. In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems.* ACM, 2014, pp. 1163–1172.

[48] Thomas Roth and Timothy Roehrs. “Insomnia: Epidemiology, Characteristics, and Consequences”. In: *Clinical Cornerstone* 5.3 (Jan. 1, 2003), pp. 5–15. ISSN: 1098-3597. DOI: 10.1016/S1098-3597(03)90031-7. URL: http://www.sciencedirect.com/science/article/pii/S1098359703900317 (visited on 06/12/2019).

[49] Jong Cheol Shin, Julia Kim, and Diana Grigsby-Toussaint. “Mobile Phone Interventions for Sleep Disorders and Sleep Quality: Systematic Review”. In: *JMIR mHealth and uHealth* 5.9 (2017), e131. DOI: 10.2196/mhealth.7244. URL: https://mhealth.jmir.org/2017/9/e131/ (visited on 04/08/2019).

[50] Børge Sivertsen. *Sleep Problems in Norway.* 2016. URL: http://www.fhi.no/en/op/hin/mental-health/sovnvansker---folkehelserapporten-2014/ (visited on 05/20/2019).

[51] Børge Sivertsen et al. "Cognitive Behavioral Therapy vs Zopiclone for Treatment of Chronic Primary Insomnia in Older Adults: A Randomized Controlled Trial". In: *JAMA* 295.24 (June 28, 2006), pp. 2851–2858. ISSN: 0098-7484. DOI: 10.1001/jama.295.24.2851. URL: https://jamanetwork.com/journals/jama/fullarticle/203083 (visited on 04/08/2019).

[52] *Sleep Better*. URL: https://www.fitbit.com/no/sleep-better (visited on 06/17/2019).

[53] *Sleepcoacher – Android Apps on Google Play*. URL: https://play.google.com/store/apps/details?id=edu.brown.sleepcoacher (visited on 06/10/2019).

[54] *SleepTown*. URL: https://sleeptown.seekrtech.com (visited on 06/04/2019).

[55] Melanie Swan. "The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery". In: *Big Data* 1.2 (June 1, 2013), pp. 85–99. ISSN: 2167-6461. DOI: 10.1089/big.2012.0002. URL: https://www.liebertpub.com/doi/full/10.1089/big.2012.0002 (visited on 06/17/2019).

[56] Giulio Tononi and Chiara Cirelli. "Sleep Function and Synaptic Homeostasis". In: *Sleep Medicine Reviews* 10.1 (Feb. 1, 2006), pp. 49–62. ISSN: 1087-0792. DOI: 10.1016/j.smrv.2005.05.002. URL: http://www.sciencedirect.com/science/article/pii/S1087079205000420 (visited on 05/29/2019).

[57] *Use Bedtime to Track Your Sleep on Your iPhone*. URL: https://support.apple.com/en-us/HT208655 (visited on 06/12/2019).

[58] Jim Waterhouse et al. "Jet Lag: Trends and Coping Strategies". In: *The Lancet* 369.9567 (Mar. 31, 2007), pp. 1117–1129. ISSN: 0140-6736. DOI: 10.1016/S0140-6736(07)60529-7. URL: http://www.sciencedirect.com/science/article/pii/S0140673607605297 (visited on 05/29/2019).

[59] Marc Wittmann et al. "Social Jetlag: Misalignment of Biological and Social Time". In: *Chronobiology International* 23.1-2 (Jan. 1, 2006), pp. 497–509. ISSN: 0742-0528. DOI: 10.1080/07420520500545979. pmid: 16687322. URL: https://doi.org/10.1080/07420520500545979 (visited on 05/29/2019).