Human Media Interaction Department
Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)

Master's Thesis

Perceived Credibility in mHealth Apps: A Case Study on a Sleep Scheduling app for Insomnia

Mohit Ahuja

Dr Randy Klaassen
Human Media Interaction Department
University of Twente

Dr Mariët Theune
Human Media Interaction Department
University of Twente

PDEng. Begum Erten Uyumaz
Industrial Design Department
TU Eindhoven/Philips

UNIVERSITY OF TWENTE.

TU/e

PHILIPS

September 24, 2018

Mohit Ahuja
Perceived Credibility in mHealth Apps: A Case Study on a Sleep Scheduling app for Insomnia
Master's Thesis, September 24, 2018
Reviewers: Dr Randy KLAASSEN and Dr Mariët THEUNE and PDEng. Begum ERTEN UYUMAZ

University of Twente
Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)
Human Media Interaction Department
P.O. Box 217
7500 AE Enschede, the Netherlands

Abstract

Insomnia is a sleep disorder faced by 10% of the global population; it affects the sufferers and the close-ones adversely. There are medicinal and non-medicinal treatments. Medicinal treatments have adverse side-effects such as daytime drowsiness, on the other hand, non-medicinal treatments don't suffer from such side-effects. The gold-standard non-medicinal treatment of insomnia is Cognitive Behavioural Therapy for Insomnia (CBT-I). It involves bringing behavioural changes in sleep and increase cognitive awareness of sleep. The number of sleep therapists providing CBT-I is limited, and waiting lines for diagnosis and treatment are long.

We solve the scalability using *Sleep Scheduler*, an Android application to provide CBT-I, this being a collaborative effort between TU/e, Philips, and Kampenhaeghe Sleep Centre. There have been attempts to digitise CBT-I, while a few of them such as Sleepio and SHUTi have been successful, a majority have failed to keep the patient adhere to their treatment. We explored the domain of persuasive design to examine the adherence of Sleep Scheduler. The role of credibility in such an app's adherence was also investigated.

During this research study, we organised three focus groups to find that people are persuaded strongly by personalisation, a persuasive design feature. *Sleep Scheduler* provides personalisation through the CBT-I based schedule it recommends. Additionally, we introduce personalisation by push notifications. By adding the participant's name in the text, the push notifications were personalised.

We did a before-after study to evaluate *Sleep Scheduler* for credibility and adherence. Using combinations of few credibility scales such as credibility expectancy questionnaire (CEQ), perceived credibility questionnaire (PCQ), and interview questions, we examined the credibility and adherence. We also recorded the user experience using user experience questionnaire (UEQ).

The app was rated highly by the participants on the credibility scales, and it was used quite regularly by our users. The role of personalisation was also crucial as the participants saw the personalised push notifications and were curious about their personal schedule. As the goal of CBT-I is to make people cognitively aware of their sleep schedules, we were successful in this project.

The participants found *Sleep Scheduler* user-friendly and motivating to use it regularly in the ten days. Thus we resolved the issues of adherence for a fixed period using a combination of personalisation and credibility parameters for *Sleep Scheduler*. The sample size for the user tests was quite small, and to make statistically prominent findings a study with a bigger and diverse crowd is the next step in our research.

Acknowledgement

I want to thank *PDEng. Begum Erten-Uyumaz* for providing the opportunity to work on this project. The past six months I worked on this project has been and will always be memorable. I believe the outcomes of this project help you in achieving your doctoral degree.

I want to thank *Dr Randy Klassen* for meeting on a regular schedule and virtually guiding me through the whole process of graduation, with simultaneously providing feedback on the project on a regular basis. The entire process would be impossible without your involvement in this project.

I would also like to thank *Dr Mariët Theune* for being the guiding light all along my stay at the University of Twente and also guiding the final phase of this project with critical evaluation for the publishing quality of the work. Her continuous communication through the second year of my EIT double degree has helped in ways I can not imagine.

I want to thank *Dr Dirk Heylen* for allowing me flexibility in the study plan and being part of the initial conversations regarding the thesis project. Along with these people, I would also like to thank people at administrative and examination departments and all my other teachers at the *University of Twente* and *KTH* during my education in *EIT Digital* program. I would also like to thank *Philips Design*, *Kempenhaeghe*, and *TU/e* for a wonderful collaborative project where I learnt a lot about working in teams and building a cutting edge product.

I want to thank my friends and family for being my support system and helping out during different phases. These friends whether they be the intern group from Philips or be it EIT classmates like *Amit Kumar Gupta*, *Jay Nagdeo*, and *Manish Thorani* and a lot of others, all of them played a small but essential role in helping me finish the thesis. My family has been fantastic as they allowed me time to grow myself away from them.

Finally, I would like to thank God for all the blessings and for giving me the enthusiasm and energy.

Contents

List of Figures

List of Tables

1 Introduction

> *For, usually and fitly, the presence of an introduction is held to imply that there is something of consequence and importance to be introduced.*
>
> **— Arthur Machen**
> (Welsh Fiction Author)

The word *insomnia* originates from two Latin words- *in-* (not) and *somnus* (sleep). This lack of (quality) sleep makes every day a hard journey for 10 to 15% of the world population suffering from this disorder [Sad10]. Clinically, insomnia is defined as the subjective perception of dissatisfaction with the quality and amount of sleep over a prolonged period, with the occurrence of dissatisfying sleep 3 nights or more per week [Rot07; Lee05; Jac+04].

The patients of insomnia are asked to record their sleeping patterns in the form of sleep diary, a paper form to record their schedules. Through a collaborative project, researchers at Philips, TU/e, and Kampenhaeghe Sleep Centre wanted to create a mobile app called *Sleep Scheduler*. The Sleep Scheduler app is a digitized version of the sleep diary. The tool should facilitate logging of sleep hours and provide non-medicinal insomnia treatments. The current project is a part of this collaborative effort.

1.1 Insomnia

1.1.1 Problem and Impact

Insomnia is formally called *Difficulty Initiating and Maintaining Sleep* (DIMS). It is defined as the inability of an individual in getting quality sleep over a prolonged period. This period can vary from a few days to weeks [Tuc]. Also, it can be a dissatisfaction with the quality of sleep during this period. Insomnia is a prevalent sleep disorder in the general population [Rot07].

Literature reveals that insomnia has a significant impact on the quality of life of individuals with insomnia [Sad10]. The co-morbidities are anxiety, depression which significantly hurt the quality of life [JRB06]. The loss of productivity and capacity for enjoyment may also result in emotional losses for the individual in private and professional life.

There is an economic impact of insomnia as well, the report *Your Guide to Healthy Sleep* estimated the loss of US$ 50 billion due to productivity decrease and US$ 16 billion due to health costs [Hea+06]. Another independent research group in Quebec, Canada found out in a cost vs benefit study that benefits of treating insomnia are far more than the economic costs involved in the treatment [Dal+09].

1.1.2 Treatments

The treatments for insomnia have great chances to bring back the normalcy from this disorder. There are pharmacologic as well as non-pharmacologic treatments for insomnia, or a combination of these two [Hea+05]. The pharmacologic solutions such as sedatives and hypnotics decrease arousal to induce sleep. These treatments have harmful side effects such as daytime drowsiness. The drug-free (non-pharmacologic) treatments don't have this disadvantage.

Through the *Sleep Scheduler* app, we provide (a part of) one of the non-pharmacologic treatments called Cognitive Behavioural Therapy (CBT). CBT for insomnia (CBT-I) is the gold-standard in drug-free solutions for insomnia as can be found in literature [Jac+04]. CBT-I is a multi-component therapy, each of its components can be provided as a single therapy as well. Four main components of it are *sleep hygiene*, *relaxation training*, *sleep restriction*, and *stimulus control* [Hor+12]. These components are explained below:

- **Sleep Hygiene**: It is a psycho-educational component, where patients are educated about sleep-friendly health practices (such as diet, exercise, and substance use) and bedroom environment settings (such as lighting, noise, temperature, and bedding). Good sleeping practices include having regular sleep routine, avoiding caffeine close to bedtime etc..
- **Relaxation Training**: Insomnia patients have high arousal rates in general, so they don't relax enough to fall in sleep. The relaxation training component is also a psycho-educational in nature. This training can yield good results in decreasing the overall (physiologic/emotional/cognitive) arousal of the patient when they want to sleep [SA12]. This includes relaxation exercises, visualization of peaceful scenery etc.
- **Sleep Restriction Therapy**: *Sleep Restriction Therapy* (SRT) was developed by Spielman et al [SST87]. This component is based on a factor called sleep efficacy (*S.E.*) which is defined as follows:

$$S.E. = \frac{\text{Total sleep time}}{\text{Time spent in bed}} \times 100$$

  The goal of this component is to enhance the S.E., and suggests patient regulate their sleep-wake cycle by limiting the time in bed to only used for sleep, thus increasing sleep drive.
  This is achieved by first reducing the time in bed, and later gradually increasing the amount of sleep on the increase of sleep efficacy to over 90%. However, there is a lower limit of 5h on the number of hours to sleep. During the restriction, patients are not allowed to nap and they have to sleep in a single stretch. There can be daytime drowsiness as a side-effect but it reduces with increase in time in sleep efficacy.

- **Stimulus Control**: The therapy *stimulus control* prescribes instructions to avoid doing activities which keeps you alert in bed. It instructs patients to get out of bed if a person is awake, this is to acknowledge the state of wakefulness. This uses conditioning according to train the person to sleep only when the stimuli inducing sleep are present, not otherwise.

In addition to these four therapy components, cognitive therapy deals with shifting from maladaptive beliefs about sleep and education about sleep-influencing factors. We discuss the instrument used in providing CBT-I in the upcoming section.

1.2 CBT-I Recording Instruments

CBT-I is a drug-free treatment for insomnia given in form of intervention therapy. In the sessions with the sleep-therapist (also called sleep coach or somnologist), the discussion is based on the patient's perception of sleep. Doctors suggest tools such as a paper-based sleep diary to record their perception. We talk about the paper-based sleep diaries and their digital alternative in the form of Sleep Scheduler which is the focus of our research study.

1.2.1 Paper-based Sleep Diary

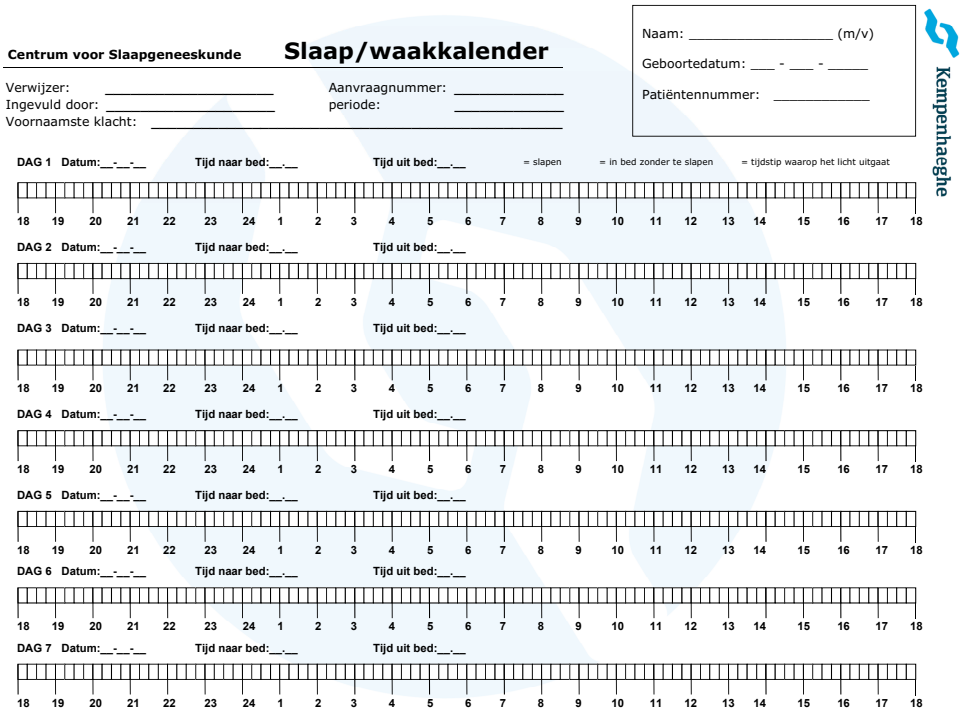
Paper-based sleep diary is a paper form to record the quantity and quality of sleep, it gives a good indication of the sleep patterns and disorders, if any, for the patient [Hea+06]. It is also helpful in the objectively assessment the sleep scheduling [Lee05], and in recommending future schedules [Ert+18]. We outline a few sleep diaries here.

1. *Kampenhaeghe Sleep Diary*: It is a paper form to be filled by the patient between interventions. The individual needs to highlight various activities using a pen. The events indicated are of time spent in bed sleeping, time spent napping, and time spent awake in bed.
One day is indicated by ninety-six 15-minute blocks as shown in Figure 1.1 from 6 pm to the next day at 6 pm. Kampenhaeghe Sleep Centre developed this diary.

2. *AASM Sleep Diary*: Also a paper form, the day is indicated as twenty-four 1-hour blocks, to shade the sleep and nap, while allowing for ways for highlighting exercise, alcohol and coffee/cola consumption, and medicine intake (see Figure 1.2). The American Association for Sleep Medicine (AASM) developed this diary.

3. *NIH Sleep Diary*: This paper form (seeFigure 1.3) is developed by National Heart, Lung, and Blood Institute (NHLBI), run by the National Institutes of Health (NIH). This diary is much more comprehensive as it allows for capturing the information of sleep during two times everyday day (first in the morning and then in the evening).

Sleep diaries in the paper form have been in use for decades and are trusted by the specialists [LFE18]. Since CBT-I has become the treatment of choice for insomnia, it is important for sleep recording instruments such as sleep diary be available ubiquitously as well.



Centrum voor Slaapgeneeskunde **Slaap/waakkalender**

Naam: _____ (m/v)
Geboortedatum: ___ - ___ - _____
Patiëntennummer: _____

Verwijzer: _____ Aanvraagnummer: _____
Ingevuld door: _____ periode: _____
Voornaamste klacht: _____

Kempenhaeghe

DAG 1 Datum:__-__-__ Tijd naar bed:__.__ Tijd uit bed:__.__ = slapen = in bed zonder te slapen = tijdstip waarop het licht uitgaat

DAG 2 Datum:__-__-__ Tijd naar bed:__.__ Tijd uit bed:__.__

DAG 3 Datum:__-__-__ Tijd naar bed:__.__ Tijd uit bed:__.__

DAG 4 Datum:__-__-__ Tijd naar bed:__.__ Tijd uit bed:__.__

DAG 5 Datum:__-__-__ Tijd naar bed:__.__ Tijd uit bed:__.__

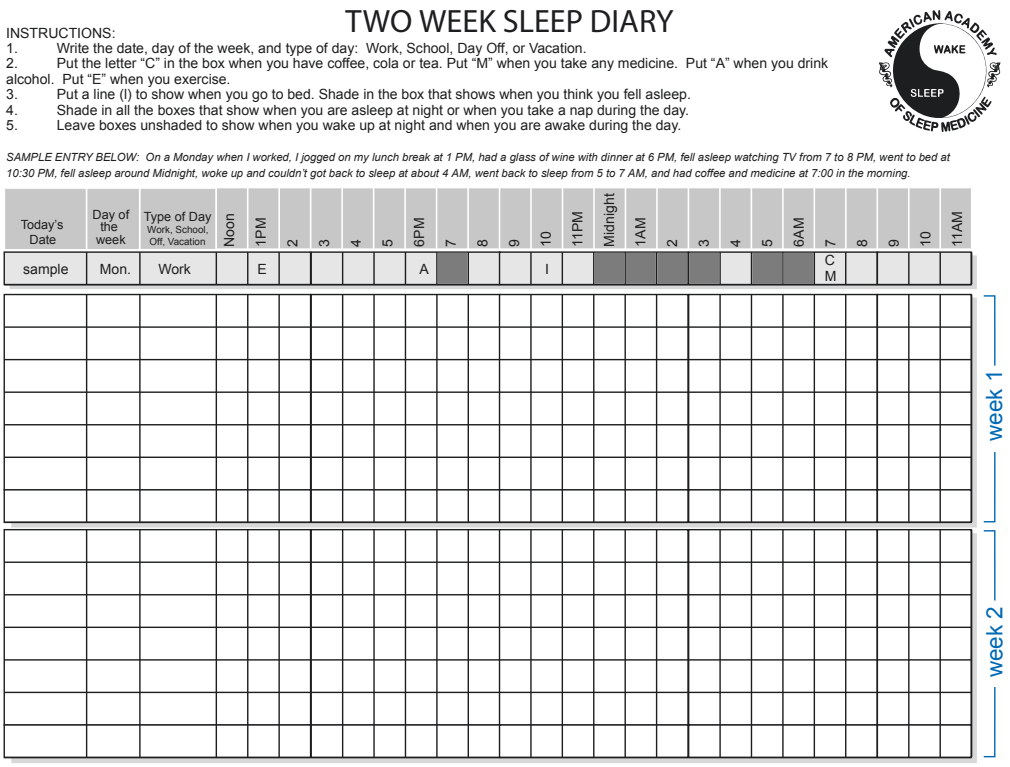
DAG 6 Datum:__-__-__ Tijd naar bed:__.__ Tijd uit bed:__.__

DAG 7 Datum:__-__-__ Tijd naar bed:__.__ Tijd uit bed:__.__

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Fig. 1.1.: Sleep Diary being used at Kempenhaeghe (complete version can be found in Appendix A)

Source: Kampenhaeghe Patiënten-slaapgeneeskunde [Kem]



TWO WEEK SLEEP DIARY

INSTRUCTIONS:
1. Write the date, day of the week, and type of day: Work, School, Day Off, or Vacation.
2. Put the letter "C" in the box when you have coffee, cola or tea. Put "M" when you take any medicine. Put "A" when you drink alcohol. Put "E" when you exercise.
3. Put a line (I) to show when you go to bed. Shade in the box that shows when you think you fell asleep.
4. Shade in all the boxes that show when you are asleep at night or when you take a nap during the day.
5. Leave boxes unshaded to show when you wake up at night and when you are awake during the day.

AMERICAN ACADEMY OF SLEEP MEDICINE — WAKE / SLEEP

SAMPLE ENTRY BELOW: On a Monday when I worked, I jogged on my lunch break at 1 PM, had a glass of wine with dinner at 6 PM, fell asleep watching TV from 7 to 8 PM, went to bed at 10:30 PM, fell asleep around Midnight, woke up and couldn't got back to sleep at about 4 AM, went back to sleep from 5 to 7 AM, and had coffee and medicine at 7:00 in the morning.

Today's Date	Day of the week	Type of Day Work, School, Off, Vacation	Noon	1PM	2	3	4	5	6PM	7	8	9	10	11PM	Midnight	1AM	2	3	4	5	6AM	7	8	9	10	11AM
sample	Mon.	Work		E					A				I									C M				

week 1

week 2

Fig. 1.2.: AASM Sleep Diary

Source: Sleepeducation.org [Ame17]

The limited number of sleep specialists limits the scalability of CBT-I dissemination. Digital forms of CBT-I covered under the umbrella term digital cognitive behavioral therapy for insomnia (dCBT-I) have come as a field to encompass the solutions which

Sample **Sleep Diary**

One of the best ways you can tell if you are getting enough good quality sleep, and whether you have signs of a sleep disorder, is by keeping a sleep diary. Use this sample diary to get started.
—Source: NHLBI

	Name							
Complete in the Morning	Today's date (include month/day/year):	**Mon***	**Tues**	**Wed**	**Thurs**	**Fri**	**Sat**	**Sun**
	Time I went to bed last night: Time I woke up this morning: No. of hours slept last night:	11 p.m. 7 a.m. 8						
	Number of awakenings and total time awake last night:	5 times 2 hours						
	How long I took to fall asleep last night:	30 mins.						
	How awake did I feel when I got up this morning? 1—Wide awake 2—Awake but a little tired 3—Sleepy	2						
Complete in the Evening	Number of caffeinated drinks (coffee, tea, cola) and time when I had them today:	1 drink at 8 p.m.						
	Number of alcoholic drinks (beer, wine, liquor) and time when I had them today:	2 drinks 9 p.m.						
	Naptimes and lengths today:	3:30 p.m. 45 mins.						
	Exercise times and lengths today:	None						
	How sleepy did I feel during the day today? 1—So sleepy had to struggle to stay awake during much of the day 2—Somewhat tired 3—Fairly alert 4—Wide awake	1						
		* This column shows example diary entries—use as a model for your own diary notes						

Fig. 1.3.: NIH Sleep Diary
Source: *Your Guide to Healthy Sleep* report [Hea+06]

are available in digital forms to provide various forms of sleep therapy [LKE17]. One such example of dCBT-I solutions is the Sleep Scheduler app, which enhances the ubiquity of sleep diary and provides scalability to CBT-I dissemination.

1.2.2 Sleep Scheduler App

The result of our study is intended to be an app, called as *Sleep Scheduler*. It is a digital version of the sleep diary, so unlike the paper forms it can be perennially available and adaptable. The features of the app finally build for the parent project will have the capability to intervene and suggest sleep-schedule to the user, thus providing CBT-I digitally. In this project we work on an earlier prototype which can act as a sleep diary

logging device and suggest sleep schedule based on Sleep Restriction Therapy, one of the components of CBT-I.

In previous user studies, it has been found that the credibility of an application has a role to play in its continuous usage and people's belief in its capability. So, this project explores questions about the credibility of Sleep Scheduler.

Apps which have been made to provide sleep therapy, have failed people to keep using them (non-adherence) [Yu+18]. On the other hand, there have been apps which people have gotten used to and they adhere to using them all the time. The field of technology for such attitude and behaviour change is called *Persuasive Technology*. These vary from content consumption apps such as YouTube, Netflix, etc; social media apps such as Facebook, Instagram, etc; utility apps like maps, email, calculator etc; health and fitness tracking apps like mySugr, Fitbit etc. There are lot of apps which have implemented features from this domain.

We will introduce some features from Persuasive Technology in our app design so people will use this app regularly as a habit. One of the models in persuasive technology, which we use primarily in this project is the persuasive system design model (PSD Model), which is explained in depth in Section 2.3.2. This model talks about the design features which can be used in a system to persuade its user for behaviour change. In total, PSD Model has 28 design features (see Figure 1.4). This study presents how we reduced the number of features for our design.

	PERSUASIVE DESIGN FEATURES			
PERSUASION CONTEXT	**PRIMARY TASK SUPPORT**	**DIALOGUE SUPPORT**	**CREDIBILITY SUPPORT**	**SOCIAL SUPPORT**
The Intent *Persuader* *Change type*	*Reduction* *Tunneling* *Tailoring* *Personalization* *Self-monitoring* *Simulation* *Rehearsal*	*Praise* *Rewards* *Reminders* *Suggestion* *Similarity* *Liking* *Social role*	*Trustworthiness* *Expertise* *Surface credibility* *Real world feel* *Authority* *Third party endorsements* *Verifiability*	*Social learning* *Social comparison* *Normative influence* *Social facilitation* *Cooperation* *Competition* *Recognition*
The Event *Use context*[a] *User context*[b] *Technology context*[c]				
The Strategy *Message* *Route*				

[a] Problem domain dependent features
[b] User dependent features e.g. goals, motivation, lifestyles, and others
[c] Technology dependent features

Fig. 1.4.: Persuasive System Design Model
Source: [LO11]

1.3 Research Setting

1.3.1 Audience

The Sleep Scheduler app would potentially impact the lives of people who have insomnia in a positive way. The audience for the finally completed instrument would be people suffering from various intensities of insomnia. However, this study is more about the user experience and perceived credibility evaluation of this application, and we will be working with healthy sleepers above the age of 18 as our participants in the studies we conduct. To check their sleep quality, participants will be given a questionnaire called the insomnia severity index to understand their current state.

1.3.2 Research Objective/Questions

The Sleep Scheduler app is meant to expedite the sleep diary logging and provided sleep restriction therapy for insomnia patients. To solve the problem of non-adherence, which is prevalent in dCBT-I solutions, we incorporate concepts from PSD Model. The challenge to solve now is to reduce the amount of features from 28 to 1. Reduction to a single feature makes it easy to design, implement, and evaluate the impact of features. We already had some interactive mock ups (in prototyping tools such as Axure RP and Adobe XD) of the same app, we will make an interactive software prototype during this project, so the inclusion of the design features is intended to be a minor redesign of existing functionality of the app.

To solve this design problem, we start by asking the design question *DQ 1*.

DQ 1 Which persuasive design feature from the PSD model (Figure 1.4) can enhance the credibility of the sleep scheduler app?

In addition to that to understand the role of credibility in persuasion and adherence to our app, we need to answer research questions *RQ 1* and *RQ 2* respectively.

RQ 1 How credible do the users find the app to facilitate managing their sleep window?

RQ 1 can be seen as combination of the following sub-questions.

SQ 1.1 What credibility scale to use to evaluate the user's choice?

SQ 1.2 How do the users evaluate the app for credibility on the chosen credibility scale?

RQ 2 As we facilitate the user in logging their sleep and wake hours, what is the adherence to this task?

The following sub-questions arise out of the research question *RQ 2*

SQ 2.1 Do people adhere to the task of logging?

SQ 2.2 What is the role of perceived credibility in adherence to the task?

Answering these questions with their rationale will set up the work of the master's thesis. During this project, we will research the state of the art literature, along with interaction guidelines (for mobile design), and implement the required features in a fully functional prototype of the app. The data collected during the study allows us to deduce the answers to these research questions. In addition to this, we want to evaluate the user experience, in general, because a good user experience can cause continuous adherence as well.

1.4 Thesis Structure

This chapter is the introduction to work done during the master thesis. The next few chapters are structured as follows:

Chapter 2: State of the Art Review
This chapter discusses the state-of-the-art work in the field of digital insomnia therapy solutions. It also presents relevant literature in the field of persuasive design.

Chapter 3: Sleep scheduler: A PSD Model Analysis
This chapter discusses how we reduced the number of persuasive design features in PSD Model (see Figure 1.4) by conducting focus groups.

Chapter 4: Application
This chapter is an elaborative explanation of the mobile application *Sleep Scheduler* which is the central element of this study.

Chapter 5: Research Methodology
This chapter discusses the methodology we followed during the research study. The core tenet of scientific studies is their repeatability, so this chapter discusses the set-up so that the study can be re-run.

Chapter 6: Results
This chapter is an presentation of the results we obtained during this research study.

Chapter 7: Discussion
This chapter is a discussion of this researh study, here we talk about the study from its implicative points and limitation, plus we discuss future work. This is where the research questions are answered in detail.

Chapter 8: Conclusion
This chapter acts as a conclusion of this research study. In this chapter, we conclude our project and summarise our findings.

State of the Art Review

2

> *State of the Art is the frenetic and relentless pursuit of doing what its best at that time!*
>
> **— Da Anunciação Marco**
> (Brazilian Author)

In this project, we design, implement and evaluate a digital CBT-I mobile application called *Sleep Scheduler*. It is an instrument to measure the quality of sleep. It has the power to intervene and suggest the hours you should sleep. The sleep diary feature is essential to get sleep information from the users.

Through this project, we plan to bring two behaviour changes to the user: (1) restricting their sleep to specified hours and (2) logging in the sleep diary. In the following sections, we will explore persuasive technology and how it can be applied to our use case. But before we dive into the sea of literature for persuasive technology, we will explore the state-of-the-art work done in the field of digital Cognitive Behavioral Therapy for Insomnia(dCBT-I).

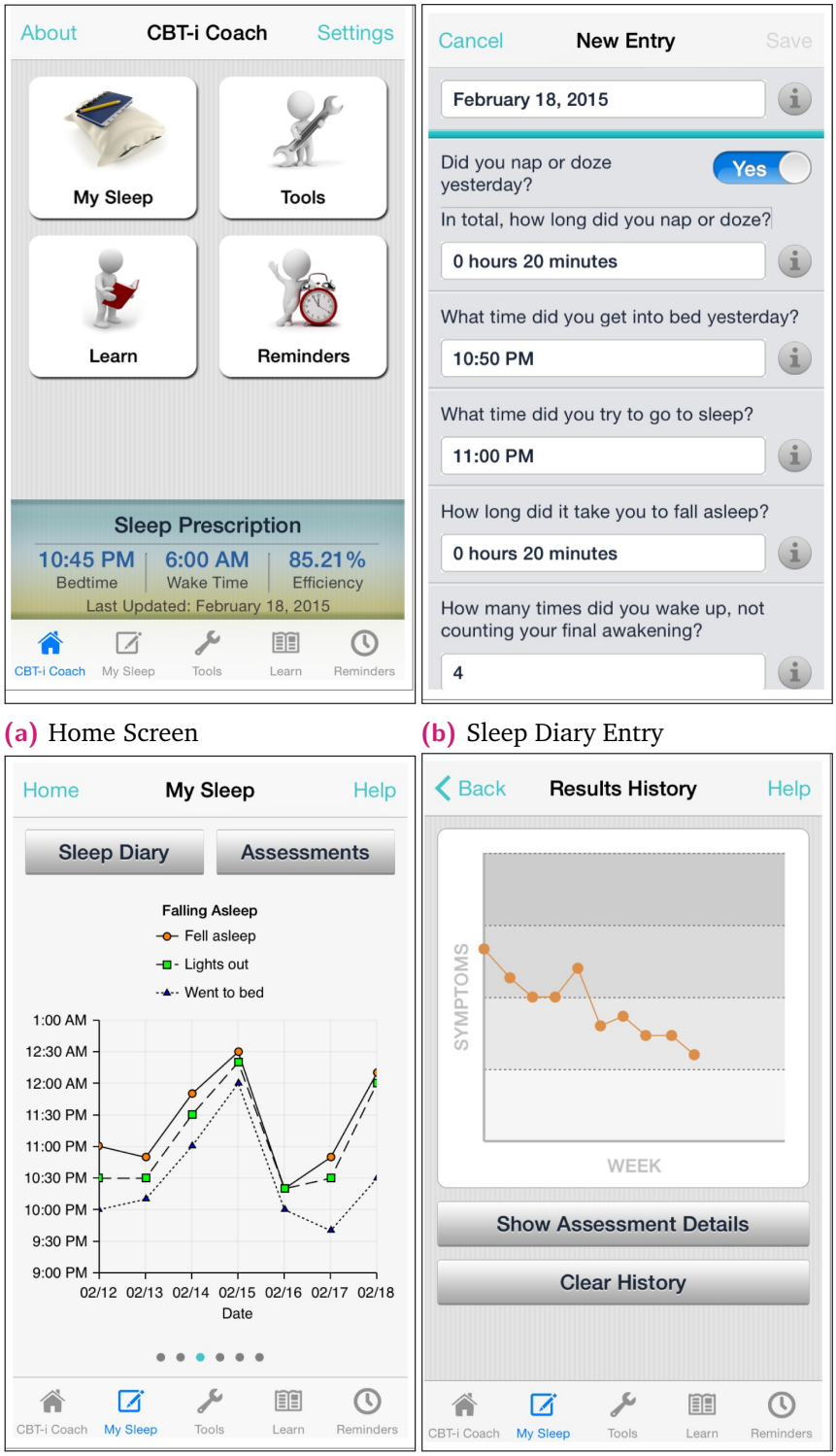
2.1 Digital CBT-I

Cognitive Behavioral Therapy for Insomnia (CBT-I) is the gold standard in the non-pharmacologic solutions for insomnia. In its traditional form, because of lack of tools and specialists, it has suffered from the problem of being unable to scale and is thus not ubiquitously available as its medicinal counterparts such as hypnotics and sedatives. The dissemination of CBT through digital means would benefit many people deprived of the diagnosis and treatment of insomnia.

Luik et al. in their review paper aggregate various digital channels through which CBT is disseminated coining the term digital CBT (dCBT) [LKE17]. They broadly categorise the dCBT provisions into three categories based on the level of automation.

1. *dCBT as support*: In this form dCBT acts as a support for conventional CBT such as just monitoring or a logging tool. In this category, the tool plays the role of a dumb terminal through which content can be provided, or data can be entered. A lot of contemporary apps for sleep management tend to do this. The role of the tool is thus limited. For example, the CBT-I coach (see Figure 2.1) is a mobile application to support face-to-face CBT [Kuh+16].

2. *Guided dCBT*: This is the most common category of digital CBT in literature [LKE17]. It can be thought of an automated program with clinical support. It



(a) Home Screen **(b)** Sleep Diary Entry

(c) Sleep Summary **(d)** Insomnia Severity Index history

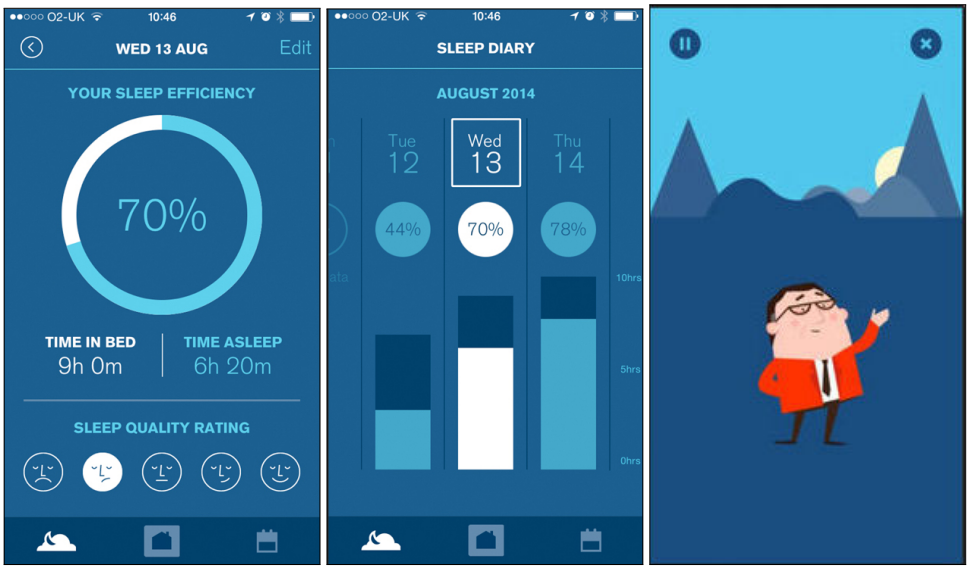
Fig. 2.1.: CBT-I Coach (iOS version)
Source: CBT-I Coach pubication in AASM [Kuh+16]

takes less time for the sleep therapist and therefore is more scalable than the conventional CBT. In this form, the digital medium renders the intervention, although clinician support is still needed.

3. *Fully automated dCBT*: This category is fully automatic and tailored dCBT, with *no* clinical support. As the dependency on sleep therapist has been removed, the

scalability is at par with pharmacotherapy. Few examples of this are *Sleepio* (see Figure 2.2) and Sleep Healthy Using the Internet or *SHUTi* (see Figure 2.3).

Sleepio is a media-rich web application also served as an iOS app. In addition to a sleep diary similar to CBT-I coach, it has a virtual sleep expert called *The Prof* which provides all components of CBT as well as other non-drug treatments for insomnia such as Imagery Relief Therapy [Esp+12].



(a) Sleep Diary Entry **(b)** Sleep Summary **(c)** (Virtual) Sleep Expert *The Prof*

Fig. 2.2.: Sleepio (iOS version)
Source: Sleepio iOS application [Esp+12]

SHUTi is a personalised and interactive web-based application designed to improve the sleep of adults with insomnia [Rit+09]. It runs an online course, where you have periodic (daily and weekly) tasks to monitor and record the subjectivity of data about your sleep. This is done by filling in a survey and a sleep diary [1].

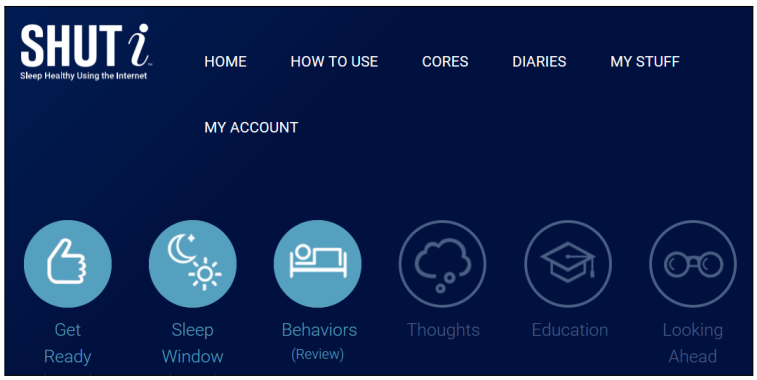


Fig. 2.3.: SHUTi (web application) interface
Source: SHUTi website (www.myshuti.com) [Rit+09]

Sleep Scheduler app is finally intended to be a Fully Automated CBT. It is to be noted that in dCBT-I solutions personalisation, contextualization, and frequent adap-

[1] As the application is behind a pay-wall, a demo can be viewed here on Vimeo.

tation are necessary prerequisites of the coaching process. Non-adherence is one of the quintessential problems in dCBT-I. The non-adherence is the rationale behind including persuasive design features in such applications [Beu+16].

In this section, we described few examples for contemporary digital solutions for insomnia and talked about various automation levels of dCBT. Through this exploration, we learnt the existing solutions and how they try to achieve dCBT-I. The upcoming section will discuss the literature of persuasive technology which we explore for potential improvements in the Sleep Scheduler app.

2.2 Persuasive Technology

Computers have affected our lives profusely. Their diffusion has presented exciting opportunities in the field of human-computer interaction (HCI). The users of these technologies no longer are just scientists or academicians, but the coverage has been covered to an average Joe. The computing technologies in the modern era are not any more just being computing devices, but, are tools to persuade users to start, stop, or continue various behaviours. Fogg defines Persuasive Technology as follows:

Definition 1. *Persuasive Technology is defined as the technology which is used to alter the attitudes and behaviour of users by persuasion and social influence, not by coercion [Fog03].*

Persuasive technologies have the power to alter behaviours for short-term and long-term [Fog09b; Fog03]. A few of the innumerable examples of its usage can be to get somebody to purchase another subscription, create more content, persist in exercising or eating healthily. Also, there can be negative examples such as to create negative propaganda or manipulations in elections. Ethics have to considered critically when designing persuasive technologies as we are influencing their behaviour.

Captology

The work on user experience in persuasive technology is rather new, from the 1990s given the inception of computers in general usage. Fogg termed the use of computers as an instrument for persuasion as Captology [Fog98]. The term is derived from the acronym CAPT, short for Computers as Persuasive Technologies. Fogg's intensive work in the field of Captology to understand persuasive role of computers has been extraordinary [Fog98; FT99; Fog02; Fog03; Fog09b; Fog09a; TF99]. An article by one of his previous students addresses him as the father of the field of Captology [Set11].

Definition 2. *Captology is the science of designing computer products (hardware/ software/ middleware) as persuasive tools [Fog03].*

Fogg's Behavioral Model

Fogg defines three factors for behaviour change: motivation, ability, and trigger [Fog09a]. The model is called Fogg's Behavioural Model (FBM). FBM asserts behaviour

is the product of motivation, ability, and triggers with each of these three factors having subcomponents. For a target behaviour, the person should be sufficiently motivated, have the ability to perform the behaviour, and be triggered to perform the behaviour. The concurrence of these components is expressed by the following equation:

$$B = m \times a \times t$$

where B = target behaviour; m = motivation, a = ability, and t = trigger (or cue) for target behaviour. Designers can use FBM for analysis and design of persuasive technologies. The understanding that trigger invocation should be done only when the user is motivated and able to perform a deed is the core tenet of this model.

Fogg discusses the basics of persuasive technology at length in his book on Persuasive Technology on the various functions that the computers, which he defines as *the functional triad* [Fog03]. The triad is the combination of three roles that computing technology can play: *tool*, *media*, and *social actor*. It becomes a *tool* when it makes some target behaviour easier to perform, do computations to persuade. It becomes *media* when simulations can be provided to persuade the user. It becomes a *social actor* when it persuades people by the principle the humans use, through positive feedback and critique, or by providing social support.

Behaviour Change Support Systems

The work by Fogg on persuasive technology is a great tool for brainstorming. It has few limitations as it creates no specific roadmap for designers to bring theory to practice. Oinas-Kukkonen has conferred various state-of-the-art works on persuasive technology, providing guidelines and models to be used for implementations. His work on Persuasive Systems Design talks about the design features and context which we as designers should understand while designing persuasive systems [OH09]. He describes the concept of Behaviour Change Support Systems (BCSS) as a key construct for research on persuasive system design, technologies, and applications [Oin10]. He defined BCSS as the following:

Definition 3. *A behaviour change support system (BCSS) is an information system designed to form, alter or reinforce attitudes, behaviours or an act of complying without using deception, coercion or inducements [Oin10].*

Kelders et al. summarised the state-of-the-art work in BCSS for health by coining the term Health Behaviour Change Support Systems (HBCSS) [Kel+16]. HBCSS are just behaviour change support system about health and wellness related behaviours. One of the examples for HBCSS is the Kristina Coaching System which is a mobile personal coaching system that measures physical activity and medication intake [Kla15]. Sleep Scheduler is also a HBCSS.

Whether we call the field as Captology or Behaviour Change Support Systems, the area has seen academic as well as industrial interest. It forms an exciting challenge as it brings various science and social science fields to be applied to technology.

In this section, we discussed persuasion technology basics and introduced you to a few basic definitions. The next section deals with persuasive technology models in depth and describes the model which fit with our project goals.

2.3 Persuasive Technology Models

This section defines two models which can be used for behaviour change: *Fogg's Behaviour Grid* and the *Persuasive System Design Model* (or PSD model, in short).

2.3.1 Fogg's Behaviour Grid

Based on Fogg's Behavioral Model (discussed in Section 2.2) [Fog09a], Fogg deduced a set of 35 behaviour change types aligned in the form of a matrix of seven rows and five columns with each row indicating the schedule of behaviour change (whether it is a one time change or change for a fixed period or is a change lasting forever) and each column indicating a type of behaviour change [Fog09b] respectively. This matrix is called Fogg's *Behaviour Grid* or simply *Behaviour Grid*. It is a tool made for designers designing persuasive technologies. We present two versions of *Behaviour Grid*, created by Fogg over time: the preliminary version and the updated version.

In both the versions, behaviour changes were discussed along two categorical dimensions: *type of behaviour change* and *schedule for the behaviour change*. The preliminary version (see Figure 2.4) had five types of behaviour change types and seven types of schedules, creating 35 type of behaviour changes. The updated version (see Figure 2.5) has condensed from 7 values in the preliminary version of the schedule to 3 values in the updated version: (1) dot (or one time), (2) span (or periodic), and (3) path (or continuous) .

In both these versions, the horizontal axis indicates the type of behaviour change, while the vertical axis indicates the schedule. The kind of behaviour change can be of five kinds, in both versions. These five kinds of behaviour changes are:
1. Performing a new behaviour,
2. Performing a familiar behaviour,
3. Increasing behaviour intensity,
4. Decreasing behaviour intensity, and
5. Stop existing behaviour.

The latest version of the grid codes these behaviour change types as colours (green, blue, purple, grey, and black).

Behaviour Grid is a great tool to brainstorm to chose behaviours we want to influence. We can combine it with Fogg's Behavioural Model (Section 2.2), to persuade for the behaviour change. The lack of a resource guide on how to implement a behaviour change or computation for ability and triggers motivates us to look at another model. The model in which explicit design features are explained is the Persuasive System Design Model suggested by Oinas-Kukkonen and Harjumaa [OH09].

On What Schedule?	What Type of Behavior Change? A: perform new behavior (unfamiliar behavior)	B: perform existing behavior (familiar behavior)	C: increase behavior (frequency, intensity, or duration)	D: decrease behavior (frequency, intensity, or duration)	E: stop behavior (cease ongoing behavior)
1 one time behavior	Take a new type of survey online	Purchase book at Amazon	Buy additional books online today	Spend less time on mySpace today	Don't eat desert tonight
2 one time behavior that leads to ongoing obligation/cost	Adopt a dog	Agree to host a party	Agree to pay more on mortgage	Pay less on a credit card one time	Refuse to continue chemotherapy
3 behavior for a period of time (X has a duration)	Play a new video game for one hour	Exercise for 30 minutes	Floss longer for two weeks.	Spend less money on books this month	Don't complain about anything today
4 behavior on a predictable schedule (X gets repeated, periodicity)	Attend online class each week for a month	Gamble online each morning at 10am	Exercise with higher heartrate each morning	Eat smaller portions for dinner.	Don't smoke after dinner each evening
5 behavior is on cue (X is cued irregularly; it's a change in habitual response)	Report any spam to AOL	Drink water at each fountain you see.	Write a longer thank you note after a dinner party	Control your frustration when driving in gridlock	Don't buy anything at checkout stand
6 behavior is at will (can perform x at any moment)	Read website privacy policy	Check computer for viruses	Check for computer viruses more often	Drink less coffee	Stop interrupting during conversations
7 behavior is always performed (X means change in habit, in way of being)	Use Google for online searching	Maintain good posture	Think thoughts of appreciation	Reduce energy consumption in home	Stop cursing

Fig. 2.4.: Preliminary version of the Behaviour Grid (shaded part indicate behaviour changes ideal for mobile persuasion)

Source: Fogg's article on The Behaviour Grid (2009) [Fog09b]

	GREEN Do new behavior	BLUE Do familiar behavior	PURPLE Increase behavior intensity	GRAY Decrease behavior intensity	BLACK Stop existing behavior
DOT One time	GREEN DOT Do a new behavior one time	BLUE DOT Do familiar behavior one time	PURPLE DOT Increase behavior one time	GRAY DOT Decrease behavior one time	BLACK DOT Stop behavior one time
SPAN Period of time	GREEN SPAN Do behavior for a period of time	BLUE SPAN Maintain behavior for a period of time	PURPLE SPAN Increase behavior for a period of time	GRAY SPAN Decrease behavior for a period of time	BLACK SPAN Stop behavior for a period of time
PATH From now on	GREEN PATH Do new behavior from now on	BLUE PATH Maintain behavior from now on	PURPLE PATH Increase behavior from now on	GRAY PATH Decrease behavior from now on	BLACK PATH Stop behavior from now on

Fig. 2.5.: Latest version of the Behaviour Grid

Source: www.behaviourgrid.org

2.3.2 Persuasive System Design Model

The Persuasive System Design Model (PSD Model) is a framework which provides a concrete way to analyse, design and evaluate the persuasion context and related techniques. There are three phases for designing a persuasive technology according to the PSD Model (see Figure 2.6). They are as follows:

1. Understanding the premises behind persuasive systems.
2. Analysing the persuasion context.
3. Design of system qualities.

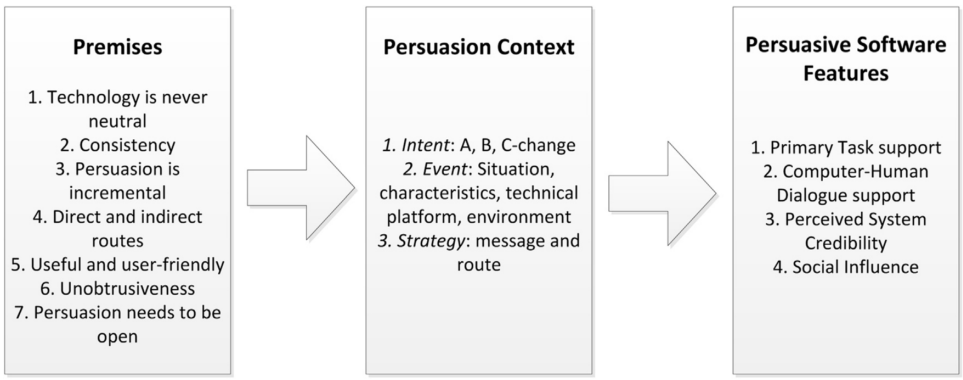


Fig. 2.6.: Three phases of Persuasive System Design Model
Source: Kelders et al. work on Health Behaviour Change Support Systems (HBCSS) [Kel+16]

The premises which need to be addressed by designers while designing a persuasive system are based on seven postulates given in [OH09], which are:

1. Technology is never neutral: it influences attitude and behaviour.
2. People like their views about the word to be organised and consistent: cognitive consistency should be a feature by default.
3. Persuasion is often incremental: all steps leading to behaviour change must be realised.
4. Persuasion strategies can be direct or indirect and depend on the ability and motivation of persuadee to process information. There are other factors such as context which may influence how the information is handled.
5. Persuasive systems must be useful and user-friendly: positive user-experience and serving needs of the user is what helps in persuasion.
6. Persuasive systems should be unobtrusive: the selection of opportune moments for persuasion is the key.
7. Persuasive designs should be open: designers should make the ideas behind and the goals of persuasion transparent.

The last two phases of persuasion context and the design features are necessary to implement a behaviour changing persuasive technology.

In the analysis of the persuasion context, a designer must look into the stakeholders, message, channel, and the larger setting in bringing the behaviour change. The individual sub-parts of the persuasion context are explained as follows:

1. The Intent: During the analysis, a designer should describe the intended outcome and the persuader's intention.
2. The Event: The analysis should comprise the final deployment environment of the final design. It should explain who is the user(s), why the use of the persuasive technology, and what form of technology do the user(s) interact with.

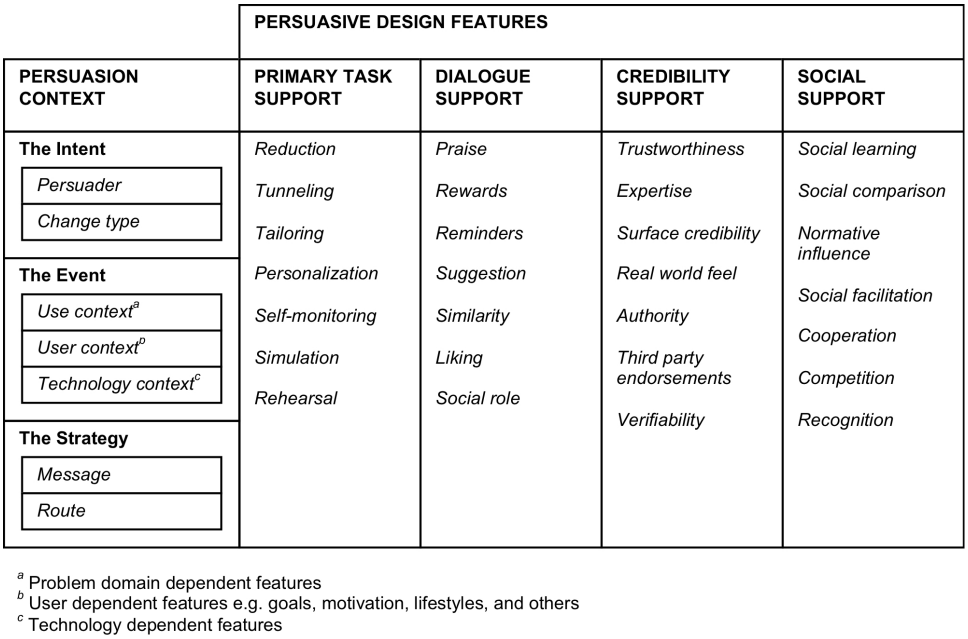


Fig. 2.7.: Persuasive System Design Model: Persuasive Context and Design Features
Source: [LO11]

3. The Strategy: The analysis for persuasive technology should include the message(s) the persuader need to convey to the user(s) and the route(s) used to achieve the behaviour change.

For implementing the behaviour change as a software tool, the PSD model describes 28 persuasive software design features. It consists of four support categories, each with seven design features (see Figure 2.7). The four categories are described as follows:

1. Primary Task Support: The design features which focus on persuasive techniques that support carrying out the target behaviour for which the behaviour change support system (BCSS) is responsible. We outline the individual features in Table 2.1.

Tab. 2.1.: Primary Task Support Feature
Source: Persuasive systems design: Key issues, process model, and system features [OH09]

Feature	Description
Reduction	Simplify the task the users are trying to do.
Tunneling	Guide the user step-by-step through a (new) process.
Tailoring	Tailor the system for the users as part of a group suiting to their potential needs, interests, personality, usage context, or other factors relevant to them.
Personalization	Customise the content or services to a user's level.
Self-monitoring	Enable them to track their progress to alter behaviour to achieve a pre-meditated outcome.
Simulation	Systems should provide simulations so they can immediately observe the link between cause and effect.

Continued on next page

Tab.2.1 – continued from previous page

Feature	Description
Rehearsal	Give ways to the end-user to rehearse their target behaviour.

2. (Computer-Human) Dialogue Support: The design features which focus on the BCSS' feedback. Individual features are described in Table 2.2.

Tab. 2.2.: Dialogue Support Feature

Feature	Description
Praise	Give (positive) feedback to the user using images, symbols, sounds, words etc.
Rewards	A system giving its user virtual (or real credits), the users will more likely achieve their goals.
Reminders	A system reminding its users of target behaviour in an ideal amount has the propensity to help its user achieve their goals.
Suggestion	Systems offering fitting suggestions based on ongoing use have greater persuasion powers.
Similarity	People look at the system, and the system should imitate them in certain ways.
Liking	The aesthetic attractiveness of a user to its audience makes it be perceived as persuasive.
Social Role	When a system takes a social role when giving feedback, the system increases its persuasive capabilities.

3. System Credibility Support: The design features which focuses on the BCSS' credibility as a persuasive system. The individual features are in Table 2.3.

Tab. 2.3.: Credibility Support Feature

Source: Persuasive systems design: Key issues, process model, and system features [OH09]

Feature	Description
Trustworthiness	A trustworthy looking system has increased capability of persuasion as well.
Expertise	A system perceived as incorporating expertise, knowledge, and competence will have increased powers of persuasion.
Surface Credibility	People make an initial assessment if the system is credible or not based on a first-hand inspection, so a system with a competent look and feel tend to be more persuasive.
Real-world Feel	A system that highlights people or organisation behind its content or services tend to have more credibility.
Authority	A system that leverages roles of authority will have enhanced powers of persuasion.
Third-party endorsements	Third-party endorsements and recommendations from other trusted entities counts.
Verifiability	The system should provide verifiable information which they claim by providing a reference where they can be cross-checked.

4. Social Support: The design features of the persuasive system which persuade the user by leveraging the social support of others. The individual elements are described in Table 2.4.

Tab. 2.4.: Social Support Feature

Source: Persuasive systems design: Key issues, process model, and system features [OH09]

Feature	Description
Social Learning	People look at each other while performing behaviours, so allow them to look at their social network who are trying to perform the same target behaviour to motivate them.
Social Comparison	System should provide users with means to compare performance with others, that will have greater motivation for them to perform the target behaviour.
Normative Influence	A system can leverage normative influence (peer pressure), thus provide means for them to see the norm, it influences their decision to increase the likelihood to adopt a target behaviour.
Social Facilitation	Users are more likely to perform target behaviour if they can find via the system that others are performing the same behaviour along with them. Thus, a system should provide the means of finding others performing the same target behaviour.
Cooperation	A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to co-operate, thus it should provide means for co-operation with others
Competition	A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to compete.
Recognition	By giving public recognition for desired target behaviour, a system can increase the likelihood of people adopting that target behaviour.

In this section, we discussed two persuasive technology models - Behaviour Grid and PSD Model in this section. The two models are great on their own, while Fogg's Behavious Grid is great as a tool for brainstorm, it is limiting in terms of providing software implementation guidelines, which is the core purpose due to which PSD model was created. We intend to use PSD model by reducing the number of design features from 28 to 1, to make it easy for implementation and evaluation. In the upcoming section, we discuss the role of credibility in persuasive technologies.

2.4 Credibility

In the book on Persuasive Technology, Fogg declares credibility as a necessary component of persuasion [Fog03]. Popular search engine Google ranks pages on their credibility when it uses its algorithm PageRank [Sha]. Although there exists no one, precise definition of credibility, scholars have attempted to define it themselves. Widespread view on this topic defined credibility to be same as believability, with even the same word for these two English terms in some languages.

Broadly, the scholars describe the term *credibility* as:

- a perceived quality [FT99; Fog03]
- made up of multiple dimensions [PGG92; SM03; FT99; Fog03; MF08]

Fogg and Tseng [FT99; TF99] attempted to describe it to comprise two dimensions-trustworthiness and expertise (see Figure 2.8).

Fig. 2.8.: Perceived Credibility as composition of its dimensions
Source: Chapter *Computers and Credibility* from *Persuasive Technology (2003)* [Fog03]

By this, they mean that a person assesses both trustworthiness and expertise to arrive at an overall credibility assessment. Trustworthiness is defined as perceived goodness and expertise is the perceived knowledge. They define that credibility is important in Human-Computer Interaction (HCI) in the following scenarios when computer products:

- act as knowledge sources
- instruct or tutor users
- act as decision aids
- report measurements
- run simulations
- report on work performed
- report about their state

The Sleep Scheduler app which is being developed can act as a knowledge source for sleep facts for sleep statistics of the user, instruct users to alter their sleep to a new schedule, serve as a decision aid to help them take a cognitive step about their sleep. Also, as it is a self-monitoring tool, it is a report about the measurements (although manually logged). Thus credibility is a valuable construct to be studied for our project.

In the book Persuasive Technology(2003), Fogg also discusses the taxonomy of credibility. There can be four kinds of credibility for a persuasive technology, and these are: (1) presumed, (2) reputed, (3) surface, and (4) earned [Fog03]. The credibility of a product can be seen as the four component playing a role in providing overall credibility.

- Presumed Credibility: The credibility which is due to the presumptions of the user about a product.
- Reputed Credibility: The credibility which is gained due to reputation in the form of citations, testimonials, referrals, or third-party endorsements.
- Surface Credibility: The credibility which the user has based on a first impression of the app.

- Earned Credibility: The credibility which is gained by a period of continuous first-hand usage.

The role of credibility is crucial in the persuasion, with it being one of the four design feature categories in the persuasive system model [OH09]. It is so critical such that an individual would not be persuaded if the source persuading them is not credible. On the other hand, gained credibility can make people anticipate features which might not even be there. We intend to explore credibility as a factor for adherence for Sleep Scheduler.

2.4.1 Credibility Studies

Scholarly interest in credibility dates back to Aristotle writing on rhetoric. Aristotle talks about three modes of persuasion to convince audiences also referred to as *artistic proofs* [Rap10; Pat].

1. *ethos*: translates to the ethical appeal (of the persuader). Ethos can be developed by choosing the right, unbiased vocabulary, right tone, and correct grammar, and syntax.
2. *pathos*: emotional appeal (of the audience). Pathos is developed by curating the delivered content by using appropriate language, contextual events and examples, and implied meanings.
3. *logos*: logical appeal (of the audience). Logos can be developed by the use of abstract language, citing facts. Proper ethos, pathos, and logos are necessary to persuade your audience.

The importance of credibility was highlighted by Aristotle's use of the term *ethos* which is source credibility. Aristotle concluded credibility is a source-based construct.

During the world wars, the *Yale Group* under the guidance of social psychologist Carl Hovland defined credibility as expertise and trustworthiness and distinguished source credibility, message credibility, and audience credulity. In contradiction to Aristotle, they suggested that credibility is a perceived quality and a receiver-based construct [MF08]. Widespread usage of mediums such as newspaper, TV gave rise to a perceived quality called media credibility. Internet and web technologies in the late 1990s gave emergence to the resurrection of the study of credibility as it conflates individual notions of the source, media, and message credibility.

Source Credibility conventionally considers features of the source: trustworthiness and expertise. In addition to that attractiveness, dynamism, for mobile apps, this can be achieved by professionalism, comprehensiveness and accuracy of the content with attractiveness and dynamism achieved by interactive features and colourful graphics. Also, an institutional-backed application such as ours which involves Philips as a stakeholder will be considered more credible than a commercial app putting irrelevant advertisements.

Message Credibility examines how content characteristics influence credibility, they can be structure, content, language, and delivery. This means the material which is doctored for the audience it is targeted for and is free of errors has the propensity to

be perceived more credible.

Media Credibility relates the various channels by which the content is delivered. The credibility of a channel (such as a newspaper) depends on the quality of information provided by it in the past. A newspaper known for its lack of research for their articles will suffer lack of credibility as well.

This subsection talked about the credibility and how it is seen as a perceived concept rather than a system's inherent property. The next subsection discusses the evaluation of credibility and the available scales.

2.4.2 Credibility Evaluation

Fogg and Tseng gave three models to represent different approaches for evaluating credibility in [FT99]:

1. binary evaluation
2. threshold evaluation
3. spectral evaluation

The three models illustrate the level of user acceptance as a function of theoretical credibility of the computer product in Figure 2.9.

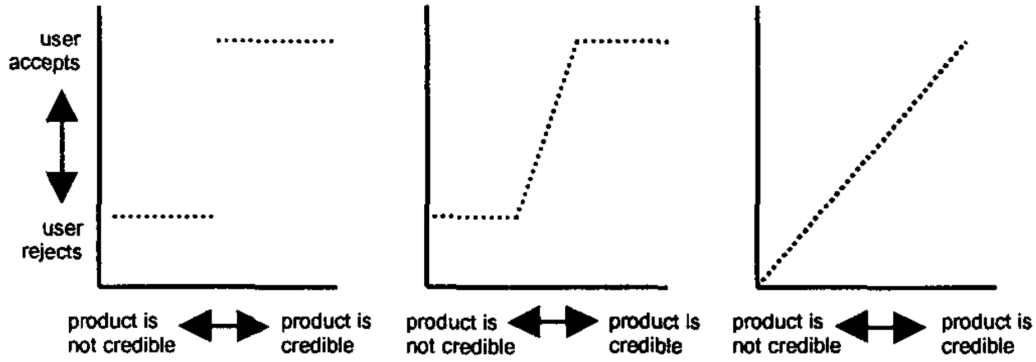


Fig. 2.9.: (From left to right) Binary, Threshold, and Spectral Credibility Evaluation Model, three different approaches to evaluating credibility.

Source: [FT99]

Binary Evaluation

Binary Evaluation is the simplest strategy where there are just two values- credible or not credible. It is adopted in case the user has - low cognitive resources to evaluate, no reference point for comparison.

Threshold Evaluation

Threshold Evaluation strategy is slightly more granular compared to Binary Evaluation, if the product has at least some threshold expectation value, only then is it deemed credible, otherwise not. Threshold Evaluation is more prominent than Binary when moderate cognitive resources or limited points of comparison are available.

Spectral Evaluation

Spectral Evaluation strategy is the most sophisticated evaluation strategy; it is in use when the user has high interest in an issue, has sufficient cognitive resources, and a considerable number of points of comparison.

All of the credibility scales for computer products fall in these three evaluation approaches. Despite its importance, credibility is hard to measure, and scales on them are tough to find. Stiff and Mongreau gives a scale of credibility with binary evaluation scales: highly credible and not credible [SM03].

2.4.3 Credibility Scales

McCroskey et al. ([MY81]) defined the source credibility scale, a popular scale in this field to calculate the source credibility. Various experiments have been done on estimating the media credibility, but lack of uniformity is an issue with such scales, with some evaluating the traditional media (newspaper, radio) to be empirically more credible than web and computer technologies, while other scales give a contradicting view of web and computer technologies being more credible. Appelman validated a scale for message credibility in 2016 [AS16].

Besides these, Ohanian developed and validated a 15-point semantic scale to evaluate both dimensions of credibility- expertise and trustworthiness. This scale has been validated psychometrically [Oha90]. This measure of credibility is similar to that used in other studies measuring the credibility of online news sources [KD06; LKC11] and e-Health websites [Cha+; CS04; VDM11] and review sites [AAL13].

To evaluate a treatment's credibility and expectancy questionnaire (CEQ), a 6-item scale divided into two sub-scales of three items each can be used [DB00]. The first sub-scale (of the first three items) measures the credibility of the treatment, and the second sub-scale (of last three items) measures the expectancy of the treatment. This scale is used to measure pre- and post-treatment differences.

Few efforts have been made to evaluate mobile apps (apart from the generic review systems in play stores) by converting existing web-credibility scales [LM16]. This evaluation study creates a questionnaire to evaluate and compare a list of existing apps amongst each other. Thus, we won't be able to use the same methodology, but we can similarly modify a few credibility scales to assess the credibility of our app. We discuss the development of credibility assessment criteria for Sleep Scheduler in Section 5.3.

2.5 Conclusion

In this chapter, we explored a few existing dCBTI solutions such as Sleepio and SHUTi. To solve adherence problems in current solutions, we explored the field of Persuasive Technology. We studied two popular persuasive technology models: Fogg's Behaviour Grid and Persuasive System Design Model.

As we implement a software prototype for the mobile application, PSD model is the natural choice for the model we chose, given its elaborative guidelines on implementing software applications with persuasive capabilities. We use the definition of credibility given by Fogg where he summarises it perceived credibility as a sum of perceived trustworthiness and perceived expertise [Fog03] (see Figure 2.8). We present more details on the PSD model's usage in our project in the next chapter and consider credibility

Finally, we talked about the research on credibility and credibility evaluation strategies with work done on evaluating computer products. The next chapter will be a discussion of the focus group which we conducted to reduce the number of design features.

3 Sleep scheduler: A PSD Model Analysis

> *Most people who are on the inside of a technology have no idea what it's like to look at from an end user's point of view. This is why they have focus groups.*
>
> — **Pat Cadigan**
> (American Author)

In the previous chapter, we explored the state-of-the-art work in persuasive system design. One of the models for designing behaviour changing systems is the PSD model (discussed in (Section 2.3.2). We have chosen this model for its simplicity and guidelines for implementation as designers. To use this model we have to understand two questions- *what is the persuasion context for the sleep scheduler app* and secondly, *which persuasive design features we should use*. The exploration of the answers to these questions is the content of this chapter.

PERSUASION CONTEXT	PERSUASIVE DESIGN FEATURES			
	PRIMARY TASK SUPPORT	**DIALOGUE SUPPORT**	**CREDIBILITY SUPPORT**	**SOCIAL SUPPORT**
The Intent *Persuader* *Change type*	*Reduction* *Tunneling* *Tailoring* *Personalization* *Self-monitoring* *Simulation* *Rehearsal*	*Praise* *Rewards* *Reminders* *Suggestion* *Similarity* *Liking* *Social role*	*Trustworthiness* *Expertise* *Surface credibility* *Real world feel* *Authority* *Third party endorsements* *Verifiability*	*Social learning* *Social comparison* *Normative influence* *Social facilitation* *Cooperation* *Competition* *Recognition*
The Event *Use context*[a] *User context*[b] *Technology context*[c]				
The Strategy *Message* *Route*				

[a] Problem domain dependent features
[b] User dependent features e.g. goals, motivation, lifestyles, and others
[c] Technology dependent features

Fig. 3.1.: Persuasive System Design Model: Persuasive Context and Design Features. The green rectangle indicates the persuasion context (discussed in Section 3.1 for the sleep scheduler), and the red rectangle indicates the persuasive design features, which we can design for (discussed in Section 3.2.2).

Source: [LO11]

3.1 Sleep Scheduler: Persuasion Context

This section describes the persuasion context (the green rectangle in Figure 3.1) for the sleep scheduler mobile application.

- Persuasion Context: Intent

  Persuader: The source of intention or the persuader(s) are endogenous, we as researchers intend to bring about the change. It is thus our utmost responsibility to design concerning user voluntariness to change.

  Change Type: The behavioural change type is temporary, as the system is intended to facilitate the users logging their sleep regularly for the ten days on which we run the user tests.
- Persuasion Context: Event

  Use Context: The use context can be determined by guidelines from Torning et al. [TO09]. The users are people suffering from mild sleep disorders or who have a problem with sleeping. The problem domain for our use context is sleep healthcare where a digital solution to sleep diary is being provided.

  User Context : This also can be determined by guidelines from Torning et al [TO09]. The specificity of the users is their difficulty in initiating and maintaining sleep. The coach and coachee in contemporary face-to-face solutions have a negotiation so that the coachee sleeps and wakes as per the suggestion of the coach. Persuasive design elements can be used to negotiate a sleep window between them and the app. It is sometimes constrained by the self-diagnosis where the patient believes that the suggested window is not going to solve their sleep difficulties. The design must take care of this constraint.

  Technology Context: The app is developed on a prevalently available mobile operating system *Android*. The negotiation will be carried out by an algorithm which was housed on the phone and to track the crashes, Philips Design infrastructure was used (by using AppCenter, a service available to Philips employees through their account).

- Persuasion Context: Strategy

  Message: The message content is going to be plain text or figures to help the patients with sleep issues.

  Route: The route for persuasion would be a combination of both argument (direct approach) and hints (indirect approach).

3.2 Sleep Scheduler: Persuasive Design Feature Selection

In the previous section, we discussed the persuasion context for the sleep scheduler. In this section, we present the methodology which we used for finding the design features we designed for. The later sections explain how we used the chosen design features.

28 design features are a lot to design and develop for in limited time as we had for this project, so to brainstorm and shortlist, we included the average user of this app

in this ideation phase. The problem to solve now was which features the users seem to be most influenced by. To find the answers to this, we did qualitative research in the form of focus groups. Focus groups are a qualitative research tool where we bring together people having similar backgrounds in a focused discussion. The organisation of the focus group has the following three steps:

1. Recruit participants for the focus group
2. Organize the discussion
3. Analyze the data obtained

3.2.1 Participants

There were 12 participants 6 males and 6 females. Three focus groups were conducted, each with four people in them. The participants were given the freedom to chose the timing for discussion they wanted to show up. The participants included young professionals and university students.

Although insomnia is a prevalent disease in the older generation (35 years and older), the bad habits are created much earlier. The irregularity in the sleep cycles late in the life is often due to the choice of lifestyles, or eating/drinking habits made when they were young. Thus the selection criteria for our focus group was healthy participants between aged between 20-35 years who spend at least 10 hours per day (on average) in front of a screen and are geographically available in Eindhoven for the focus groups.

Additionally, we needed an ethics committee approval for the participants from the university (University of Twente) for involving humans in the process. We did not go with people with sleep problems as the total time on this project was limited, and the medical permissions required for adding that inclusion criterion would have extended the project timeline. We invited the participants in three different ways described as following:

1. Public event on Eventbrite[1]: This was done on a public event publishing website called Eventbrite. We created an event for a group discussion while learning about persuasive system design. The participants were allowed to chose the time for the discussion.
2. Sharing on social media applications such as Facebook and WhatsApp. We shared the event details of Facebook groups for expats in Eindhoven and WhatsApp groups I was a part of, with a requisition to share along the network of friends. This was the most efficient method to reach as was indicated by the organic outreach of the event details.
3. Putting posters on university (TU/e) boards. The poster had posted on the TU/e public display boards to invite participants can be found in Appendix C. The QR code was a link to the Eventbrite page as mentioned in item 1.

[1]The event can be visited at https://habitformers.eventbrite.com (a screenshot can be found in Appendix B).

The only prior knowledge that was needed from the participants was the ability to read and speak English to understand each other for the discussion and the consent form.

3.2.2 Focus Group Setup

The participants we chose were from different backgrounds, and many of them were not aware of persuasive design as their background. This resulted in the focus group discussion being organised in the form of a workshop. In the first half, I presented the persuasive system design and its capabilities with some examples. In the second half, the participants were given a mock problem to solve by discussing amongst them, my role being the moderator as well as the observer in the focus group.

Persuasive Design Presentation

A presentation on persuasive design model was given by me where I explained the different persuasive design feature with simple explanations and examples. To help with the introduction and as a take away from the workshop, the participants were given a web app[2] explaining each of the features. A small break for 2-5 minutes followed the presentation.

Focus Group Discussion

The presentation was followed by a focus group discussion with a mock problem. The reason for giving them the mock problem was to understand the rationale the participants will use to alter their behaviour (or of someone they know closely) towards sleep.

The mock problem was to think of possible mobile application features that could influence sleeping them to a better schedule. The participants were suggested that they could use the tools talked about in the presentation if they wanted to, but were given freedom to chose if any other solution methodology they feel was a good fit as well.

To steer the discussion smoothly avoiding any digressions or lull, we set these questions as the discussion points.

1. Which features do people choose to use the solution they believe is a solution for sleep application?
2. The reason why they chose this feature?
3. What is the perceived advantage of the solution they come up with?
4. What is the perceived disadvantage of the solution they rejected?

The data was recorded in the form of audio and transcribed, and analysed. During this phase, no other data was collected from the user.

[2]Available at https://ahujamoh.github.io/flashcards-vue/ (a snapshot can be found in Appendix D).

3.3 Data Analysis

3.3.1 Thematic Analysis: Process

The data collected from the focus group was analysed using thematic analysis, a well-known analysis method for qualitative data created by Braun and Clarke [BC06]. This analysis is generally used in the field of psychology to find themes or patterns. It is a 6 phase process indicated by Figure 3.2. The six stages are explained succinctly here:

1. Familiarize yourself with data: This stage includes the transcription of data and multiple re-reads of data, so the researcher has an initial grasp of the data.
2. Generate initial codes: Coding interesting features in the data set systematically. This can be based on answering a research question, or research question can be a result of this process.
3. Search for themes: Collating the codes into themes/feature sets/patterns.
4. Review themes: Refinement of candidate themes creating out of collation in the previous stage.
5. Define and Name Themes: Another level of refinement, by giving a name based on what story it tells. In this stage, we generate clear definitions for each theme.
6. Produce the Report: Final stage of thematic analysis, relating to the literature and research question, with the help of compelling examples, produce a scholarly report of the analysis.

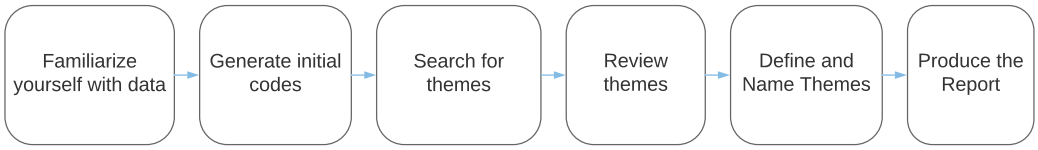


Fig. 3.2.: Thematic Analysis process as explained by Braun and Clarke [BC06]

To highlight the themes Nvivo 12, a famous qualitative data analysis tool [Ltd18] was used. For our analysis, the transcription was coded using pre-decided themes, each theme indicating a design feature of the PSD Model [OH09]. We wanted to understand the overbearing theme during the conversation. To validate our findings, another researcher also coded the project, and we validated using intercoder reliability.

Another way to analyse qualitative data was given by Stewart et al. which is content analysis for focus groups in [SSR07] which were considered for analysis, but we proceeded with thematic analysis due to experience in our team.

3.3.2 Thematic Analysis: Results

The thematic analysis are presented in Table 3.1 partially (the complete analysis can be found in Appendix I). The coverage percentage indicates the amount of text (in transcription) which was coded as belonging to a specific sub-theme.

Tab. 3.1.: Thematic Analysis of Focus Group Discussions (only theme, subthemes, and coverage)

Theme	SubTheme (Coverage)
Primary Task	Personalization (13.89%) Reduction (6.15%) Simulation (5.14%) Tailoring (4.05%) Tunneling (1.94%) Self-monitoring (1.34%)
Dialogue	Liking (6.48%) Similarity (4.93%) Rewards (4.84%) Suggestion (3.17%) Praise (2.31%) Social Roles (1.59%) Reminders (1.45%)
Credibility	Trustworthiness (2.96%) Authority (2.18%) Surface Credibility (2.12%) Expertise (1.76%) Verifiability (1.03%) Third Party Endorsement (0.55%)
Social	Cooperation (9.23%) Normative Influence (3.75%) Social Facilitation (2.76%) Social Learning (1.68%) Competition (0.67%) Social Comparison (0.37%)

We organised the focus group to find the answer to item *DQ 1*: which persuasive design influences people the most? The results were on an expected note, with *Personalization* coming as the most prominent feature on analysis of focus group discussion. A few quotes which indicate the importance of the personalisation are presented here:

You can receive a remark do you want to do it this way, or you can do it in a better way

If you can personalise the app. Then everybody can choose what puts them to sleep. You can't have a standard way like music put everybody to sleep. Maybe it puts me to sleep doesn't put them to sleep. It might wake somebody else up.

I would prefer to have personalised advice.

In the past, researchers have also found Personalisation to be a critical parameter for persuasion. Peng et al. highlight the value of personalisation in adherence of mHealth apps [Pen+16], the study involves reviewing user perception of mHealth apps. Lehto et al. also consider personalisation as one of the core design features for

accomplishing the primary task [LO11]. Kelders in her paper on HBCSS also discusses personalisation for targeting an HBCSS [Kel+16].

Personalization can be as minimum as just mentioning the name [Dij05] to as maximum as the adaptation (implicitly adjusting to user's usage) or the adaptability (explicit control is given to the user) of a system. Our app already adapts to the user's sleep based on the algorithm it uses to predict the sleep schedule. To keep the changes minimal, we used the approach where in addition to the adaptation (by previous design), we provided the personalisation through push notifications by introducing the name of the participant in the push notifications.

3.4 Conclusion

We discussed the persuasion context of Sleep Scheduler followed by the process of focus groups we used to find the persuasive design feature we should design for. The analysis of the focus group data was done using thematic analysis, and we figured out that Personalisation was the most influential design feature. The design for our app applies personalisation in one way: adapting to the user's sleeping schedule, due to the algorithm we use. Additionally, we personalise the app by adding the name of the participant to push notifications.

In the next chapter, we address the Sleep Scheduler app's design process and implementation details (albeit on a higher level). The upcoming chapters will explain the role of push notifications in our application in detail as well.

4 Application

> *In mobile, people love having single-use case experiences. They want low friction to getting to the application's use case.*
>
> — **Dave Morin**
> (Founder, Path Inc.)

The chapters till now introduced the problem we are trying to solve, the state-of-the-art work behind it, and then the analysis of the sleep scheduler app according to the PSD model. This chapter describes the app Sleep Scheduler in detail. We start by discussing the background of this app, the design process we followed, and finally, we talk about the implementation.

4.1 Background

People who have insomnia have been asked to record their sleep. There are wearables in the market which can be used for this, but these wearables do not capture the subjective feeling about the quality of sleep. Sleep Scheduler is used to track the quality of sleep by their perception. It is a digital version of Kampenhaghe sleep diary (described in Section 1.2.1). The app offers a sleep diary to record sleep, awake in bed, and napping hours, but in the future can record the moods felt about the sleep and also thoughts about sleep during daytime.

Researchers on the overarching project had worked with sleep therapists and patients to understand the interaction needed for the app for its usage, a few of their works can be found in [Uyu+17; Ert+18]. The application Sleep Scheduler has been tested in form of interactive mockups and this project is its conversion into a working software prototype with persuasive desgin features, we will discuss the app in more details in the upcoming sections.

4.2 Design Process

The Figure 4.1 illustrates the design process. We start by creating personas, write user stories, convert them into system requirements for Sleep Scheduler. Then we include users in the ideation phase in evolving the requirements, create the prototype, and test it with users for a few days. We went from version 1 (which was a functional prototype in Azure RP, already created at the beginning of this project) to version 2 (a functional prototype in Adobe XD, created during this project) to a final prototype on Android. These iterations helped us evolve the versions.

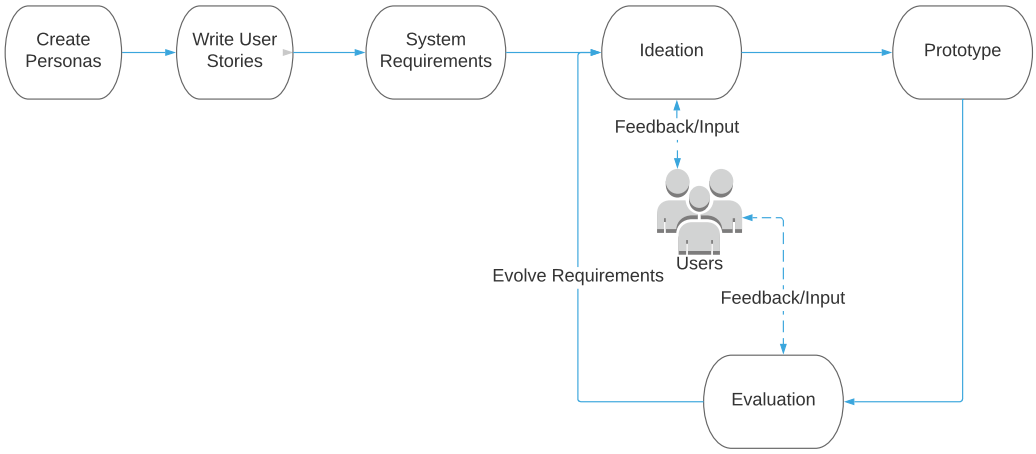


Fig. 4.1.: Design Process followed for creating Sleep Scheduler

4.2.1 Personas

To better understand the use case in which our app is created. We thought of two personas: (1) John Smith, 27-year-old young professional and (2) Jena Smith, a 24-year-old university student.

John has a hectic lifestyle, where his work is his foremost priority, and the sleep schedule generally on the weekdays is okay but mostly on weekends, the late night parties he goes leave an impression on his Mondays. He has started taking naps in the office.

Jena has made significant strides in academics, but lengthy assignments, plus playing mobile games with friends late in the night means she sleeps very less (around 5 hours) on weekdays and very long on weekends. Thus her sleep schedule is erratic.

They both have the habit of prioritising their parties and work over good sleep, and none of them has a regular sleep schedule. As they grow up, this will only become more irregular, and they need an app like Sleep Scheduler to make them aware of their sleep schedules. They know their sleep schedules are out of order and it has started affecting their work and study lives respectively.

4.2.2 User Stories

These user stories illustrate their use of the app.

1. John/Jena should be able to log their *sleep*, *awake in bed*, and *nap* hours.
2. John/Jena should be able to see a summary for their sleep for the last seven days and see the recommended sleep window shift as recommended by the system.
3. John/Jena should be able to negotiate a sleep schedule to facilitate sleep on a better sleep schedule. Thus *sleep restriction therapy* principles should be used.
4. John/Jena should be able to see their negotiated sleep schedule.
5. John/Jena should be able to see all the sleep diaries they have logged.
6. John/Jena should be able to delete the diaries they have already created.

Additionally, by our focus group research (chapter 3), we know that personalisation is one of the critical persuasive design features. Researchers have used push notifications to influence the compliance of self-reporting sleep disturbance data for breast cancer patients receiving chemotherapy [Min+14]. On the lines of [Dij05] where they use the name of the participant to personalise the experience, we will use only the name of the participant for personalising the push notifications. The sleep schedule prediction algorithm takes care of the personalisation through adaptation (by personalising the recommended schedule).

The only other form of personalisation (from the PSD model) which we add to the existing design during this project are the push notifications. Although it seems too limiting in terms of personalisation as a feature, we believe it is not as the app now gives personalisation in two forms implicitly and explicitly. Implicitly in the form of adapting to the user's sleep schedule to find their personalised schedule and explicitly in the way of these personalised push notifications.

We outline that this is the only way we explicitly include the results of our focus groups in the app. We intended to keep the design changes minimal in an iteration of this process, thus quickening the process of prototyping and evaluation.

7. John/Jena should receive personalised push notifications, at some intervals.

Additionally, the involvement of the sleep therapists and the literature of sleep restriction puts some more user stories in the app.

8. John/Jena should be able to negotiate the sleep only for two times or less; instead of negotiating the third time, they should see a message to try the recommended sleep hours for next three days. (This comes as a recommendation of sleep therapists.)
9. John/Jena should be given the sleep restriction therapy, and thus should not be recommended to sleep less than 5 hours. (This is from the literature on Sleep Restriction Therapy.)

The application should thus have the minimum capability for the following:

1. Sleep Scheduler should have a way to indicate the user's daily sleep log.
2. Sleep Scheduler should have a way to allow the user to see their sleep logs.
3. Sleep Scheduler should be able to give user capability to add their sleep logs.
4. Sleep Scheduler should be able to calculate the best sleep window for the user, based on their past sleep.
5. Sleep Scheduler should be able to indicate the best sleep window to the user.
6. Sleep Scheduler should allow for a negotiation of sleep hours to provide sleep restriction therapy.
7. Sleep Scheduler should be able to handle push notifications.
8. Sleep Scheduler should only allow the negotiation to be done only twice, suggesting user to sleep on recommended sleep hours.
9. Sleep Scheduler should give sleep restriction therapy by limiting to the minimum 5 hours of sleep.

These requirements can also be visualised in the form of the use case diagram in Figure 4.2 for sleep scheduler. There are two types of actors involved: the app user and the (push) notifier. The (push) notifier in the real-world scenario can be a bot or a therapist, but in this case, I was the person who used the Wizard of Oz method to personalise push notifications.

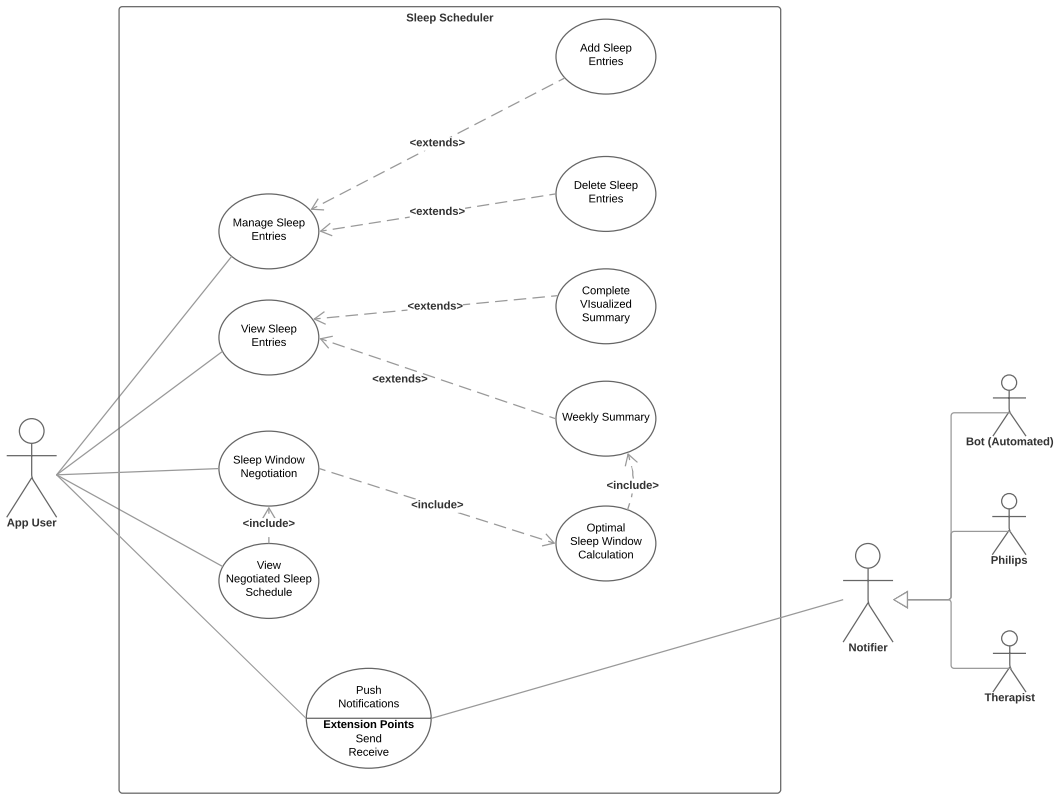


Fig. 4.2.: Use Case Diagram from Sleep Scheduler App Android Prototype

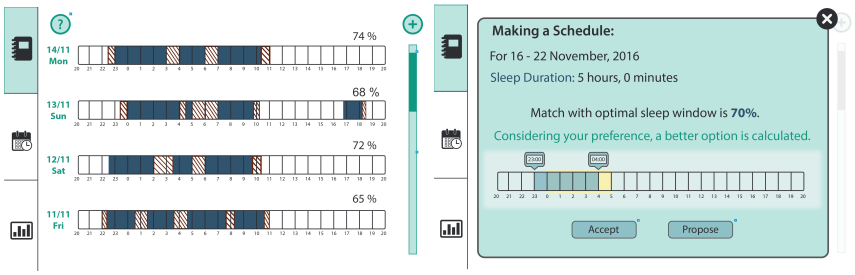
4.2.3 Iterations

The prototype evolution for Sleep Scheduler Application went through the three stages of ideation, prototype, and evaluation. During the ideation stage, the sleep therapist or the user were involved in each of the iterations to take the feedback or the input for brainstorming for the prototype. During the prototype phases, the researchers on this project worked and did internal testing of prototypes. During the evaluation stages, end users or the stakeholders in the form of sleep therapists were involved in taking feedback about the prototypes. These feedbacks were later converted to inputs for the next stage.

In the Figure 4.3, we show the prototypes for this app that were created in earlier phases and evolved before the start of this research study. During this research study, we evolved from version 1 (Figure 4.3) to version 2 (Figure 4.4) to the final android prototype (Figure 4.5).

Version 1

The prototype went from being low fidelity prototype to a high fidelity prototype in the prototyping tool called Axure RP[1] (see Figure 4.3). It was tested with real users, their feedback was noted, and I joined the project at this point. The results were used to generate requirements for the next iteration.



(a) Sleep Diary Main Screen

(b) Negotiation Feature based on Best Fitting Sleep Window

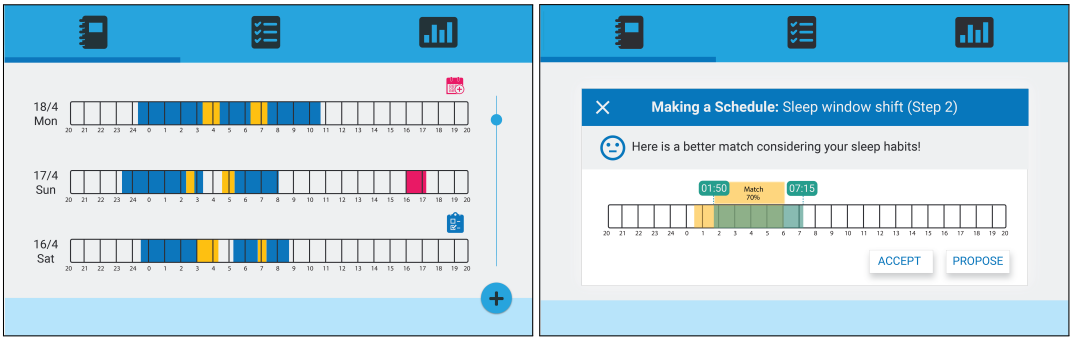
Fig. 4.3.: Screenshots from Sleep Scheduler App prototype version 1[2]

Iteration 1: Moving from Version 1 to Version 2

The previous feedback was helpful in iterating from version 1 to version 2. The interface was polished following material design guidelines of Android apps as we planned to develop the prototype as an Android app [Goo18]. The version 1 was created for a tablet screen, while version 2 was designed for a mobile screen.

There was a visual overhaul from the previous prototype (see Figure 4.4) with changing colour choices, adding material design icons, and shadows, which were not in the earlier version. Another researcher on the same project tested this version's usability.

This prototype was created using Adobe Illustrator to create the screens and interactions were simulated using Adobe XD[3], another tool for making interactive prototypes similar to Axure RP. Another researcher working on the same project evaluated the interface.



(a) Sleep Diary Main Screen

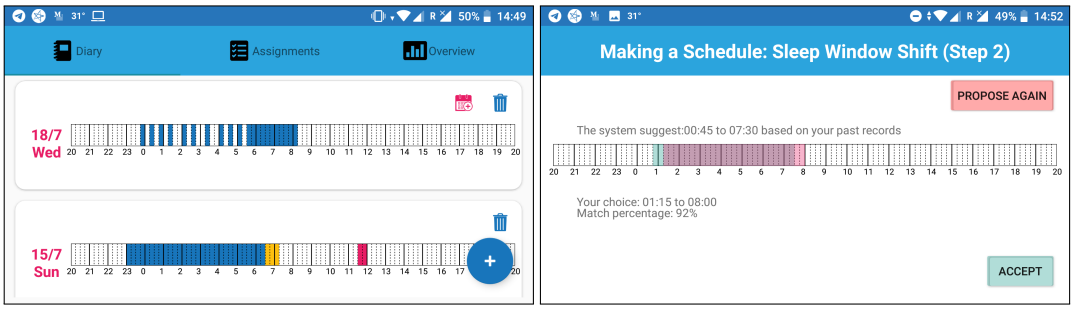
(b) Negotiation Feature based on Best Fitting Sleep Window

Fig. 4.4.: Screenshots from Sleep Scheduler App version 2[4]

[1]The tool can be found at https://www.axure.com
[2]This prototype can be made available on request to the primary researcher of this project.
[3]The tool can be found at https://www.adobe.com.

Iteration 1: Moving from Version 2 to Android Prototype

The Android prototype (see Figure 4.5) was very close to the version 2 (see Figure 4.4), with only minor changes to adhere to mobile phone. This was a functional prototype with functioning input handling. The requirements were generated from the feedback we received from the evaluation for prototype version 2 and the results of focus groups. The users wanted an application to interact with which personalises for them and looks modern (we included material design guidelines to integrate them).



(a) Sleep Diary Main Screen

(b) Negotiation Feature based on Best Fitting Sleep Window

Fig. 4.5.: Screenshots from Sleep Scheduler App Android Prototype

The system was built with a capability to handle use cases indicated in the use case diagram Figure 4.2.
- for storing and retrieving sleep entries and summary.
- to visualise sleep summaries.
- to calculate the best fitting sleep window based on the algorithm described in [Ert+18].
- to allow app user to negotiate sleep window and later allow them to view it.
- to report crashes.
- To handle push notifications: this was the way we added personalisation, which we got a result from the focus group

Additionally, we made it sure that the sleep data was never published out of the app, thus creating a closed system within the app for the personal data of the user. The push notification was personalised only to include the name so that it appears to the participant that they have been specially sent a message from the app. The push notifications were generic enough not to need the data from the participant for this iteration.

The key differences between the version v2 and final Android prototype were as follows.

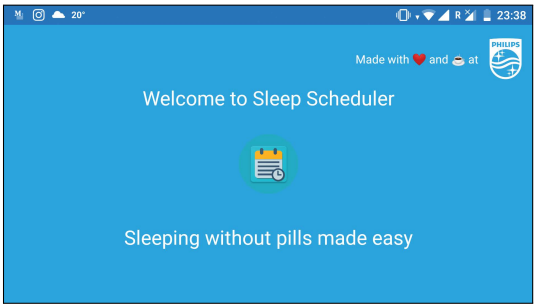
1. A launch screen with Philips icon: The Philips icon was placed on right top in Figure 4.6a.
2. Real addition and deletion of sleep logs, showing summary. The home screen of Sleep schedule is Figure 4.6b. The deletion was possible by pressing trash icon (placed on right top) on sleep diary record, to add a new diary, the (+) button is pressed. This launches the screen Figure 4.6c which can be filled using touch interaction, and options were given as buttons on top.

[4]This prototype can be made available on request to the primary researcher of this project.

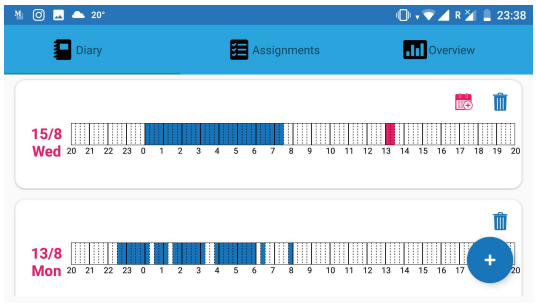
The four buttons on top indicate asleep, awake in bed, nap, and awake respectively from left to right. A full diary looks like Figure 4.6d, with blue indicating sleep, yellow indicating awake in bed, red indicating nap, white meaning awake (outside bed).
On pressing **SAVE**, the new diary is added. By default, the diary is filled with white colour which implies the participant is awake all the time. Thus the awake button can be used to erase wrongly added sleep interval in a sleep diary.

3. Implemented best sleep window shift based on sleep negotiation algorithm given in [Ert+18]. This algorithm predicts an ideal sleep window called as Best Fitting Sleep Window.
4. Negotiation interaction added based on Best Fitting sleep window shift.
5. Handling push notifications.

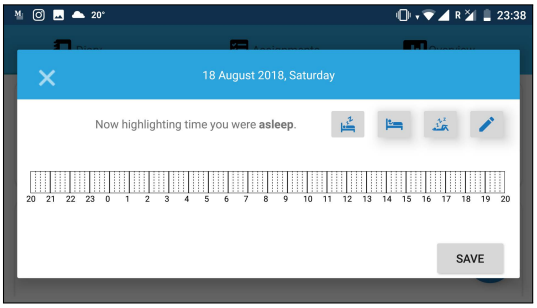
We discuss the last three differences in the upcoming section (Section 4.3.1).



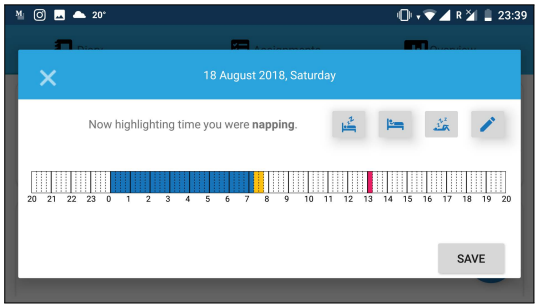
(a) Sleep Scheduler Launch Screen



(b) Sleep Scheduler Home Screen



(c) Sleep Scheduler Addition Screen (Empty)



(d) Sleep Scheduler Addition Screen (Filled)

Fig. 4.6.: Launch Screen, Home Screen, Addition Screen for Android Prototype

4.3 Implementation

4.3.1 Implementation Details

Best Fitting Sleep Window

To calculate the recommended sleep hours, we use the algorithm given by Begum et al. in [Ert+18] where they use the term *Best Fitting Window Shift*. Best Fitting Window Shift is the optimal sleep schedule, taking into account the last one week (or last seven sleep diary entries) of data to predict the best schedule for sleep hours which fits the current sleep schedule of the participant.

It is to be noted that the algorithm proposes a schedule where the sleeper should have a single sleep window and no naps (even when the original schedule may have naps) on the lines of Sleep Restriction Therapy (described in Section 1.1.2).

Negotiation Interaction

Here is how the negotiation interaction takes place.
1. Once the user has filled the entries for seven days, they are asked to make a schedule for their sleep (see Figure 4.7a).
2. On pressing **NOW**, They are shown the previous 7-day summary and told their ideal sleep hours (as text) based on Best Fitting Sleep Window (see Figure 4.7b).
3. On pressing **>** button, they are asked to set up duration for sleep (this duration is computed using the algorithm in [Ert+18], allowing the user to chose between $\pm$ 15 minutes of the existing duration (see Figure 4.7c).
4. On pressing **PROCEED**, they are shown a screen to choose their sleep window (see Figure 4.7d).
5. On pressing **SUBMIT**, the recalculated schedule is shown (it is the average of ideal sleep window and chosen sleep window).

Figure 4.7e indicates the recalculated schedule (**PROPOSE AGAIN** button is enabled). If they want to repropose the choice of the window, they can press the **PROPOSE AGAIN**.

There is a limitation on only allowing schedule proposal two times, based on the recommendation of the sleep specialists. If two proposals are made Figure 4.7f (with disabled **PROPOSE AGAIN** button) is the screen which is presented.
6. The final schedule is then presented on presenting **ACCEPT**.
7. The assignment tab is also updated (see Figure 4.7h) with the recommended schedule. It is to be noted that if a participant does two proposals, then they are asked to follow the schedule for three days, otherwise for seven days.

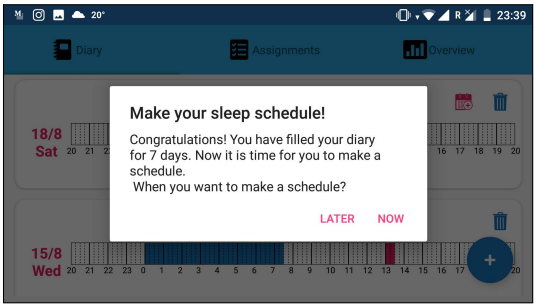
The process has been presented in Appendix E.

Push Notifications

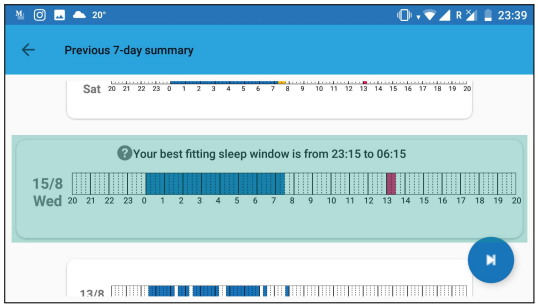
To test for the personalised notifications made up by me as the project owner, several messages were tested internally with friends and family. I typed in the notifications on a text editor and asked few friends to rate them, giving them the information that the data cannot be fetched from the device for the participant. These text were brainstormed by the researchers on this project. The final list I received were these five messages. <Paricipant Name> was a placeholder for actual participant's name.

- Hi <Participant Name>, Did you log your sleep and negotiated the future schedule? It is almost a week.
- Hi <Participant Name>, Hope you had a nice sleep, fill those hours in the sleep scheduler app soon if you haven't already.
- Welcome to our user test. :)
- Hey <Participant Name>, Did you sleep enough Saturday night? National Sleep Foundation[5] recommends 7 to 9 hours of sleep for your age group. Don't forget to make an entry in the sleep scheduler.

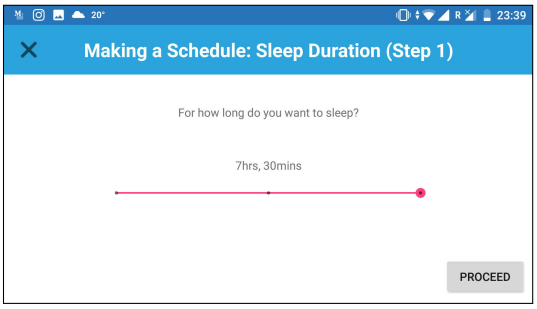
[5]https://sleepfoundation.org/



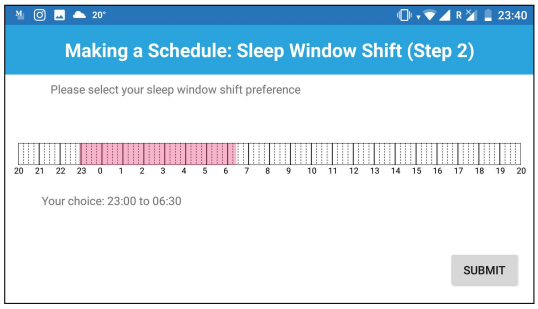
(a) Invitation to make schedule after 7 days



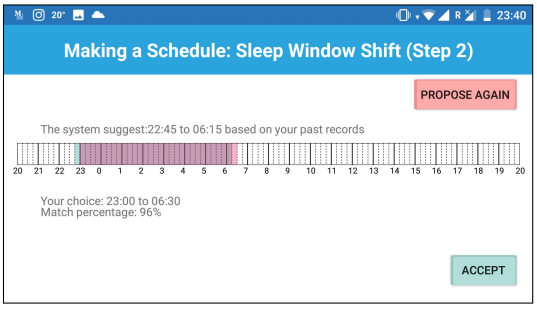
(b) Summary (in the list) and Best Fitting Sleep Window(in text)



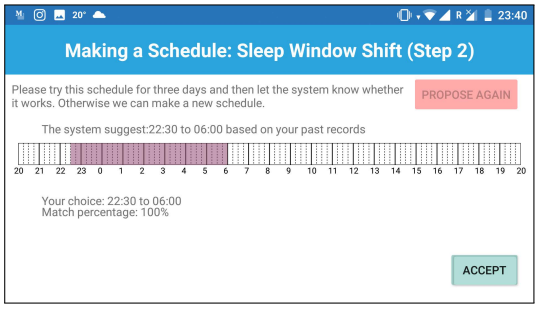
(c) Choosing sleep duration



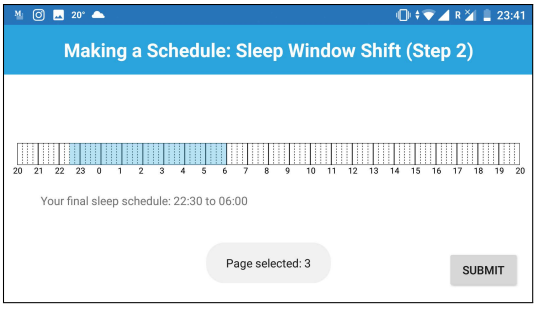
(d) Choosing sleep window



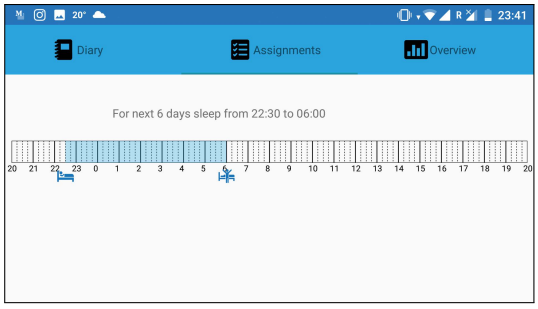
(e) Recalculated schedule before 2 proposals (enabled PROPOSE AGAIN button)



(f) Recalculated schedule after 2 proposals (disabled PROPOSE AGAIN button)



(g) Finally Negotiated Schedule indicating the accepted recalculated schedule



(h) Updated Assignments Tab indicating the finally negotitated schedule

Fig. 4.7.: Sleep Scheduler Negotiation Interaction screenshots

- A sleep tip, just for you, <Participant Name>. Your body needs time to shift into sleep mode, so spend the last hour before bed doing a calming activity such as reading. Source: National Sleep Foundation[6]

To send the push notifications a Microsoft tool called App Center. This tool was used as it provides a researcher to change the content/schedule of the push notification

[6]https://sleepfoundation.org/

on-the-fly. Also, Philips could also replace this with their push server in the future. For implementing the push notifications, we had to include a library which could handle push notification on the device and send the crash logs, on the device. The sleep data was never pulled out, and the crash stack trace if pulled, was without the sleep data attached to it .Figure 4.8 indicates how the system was set up for sending the push notifications and receiving crash logs. Push notifications were data pushed by me, but could easily be automated; only crash stack trace pulled if there was a crash (the app was tested internally to prevent any crashes). The sleep data was NEVER pulled out of the app, neither during the experiment nor after it.

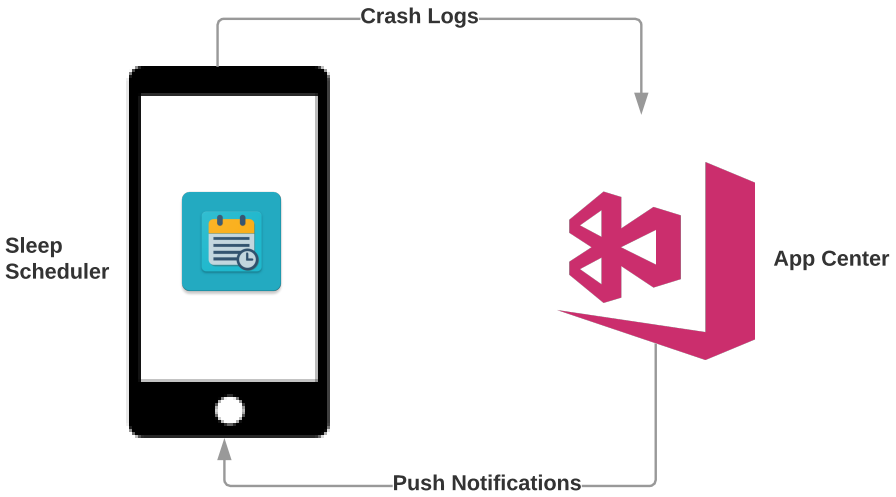


Fig. 4.8.: To send the push notifications we set up the AppCenter to send the notifications and receive crash reports; the app was set up to receive push notifications and send crash reports

An example push notification is given in the Figure 4.9. More details of push notifications (when they were sent) are discussed in the next chapter.

4.3.2 Application Architecture

This section will discuss application architecture details for Sleep Scheduler. By using the architecture, we achieved an efficient separation of concerns in the software design.

To create our android application, we followed a popular architecture pattern called Model-View-Presenter (see Figure 4.10). There are three layers in it: Model, View, and Presenter.

The three layers are described as follows:

- **Model**: This layer is an abstraction of the data in the system. It is responsible for managing databases, communicating with APIs, and all exchange of data. In our system, we have two types of data which should be stored and retrieved - one of the sleep and the other of negotiation.
- **View**: This layer is an abstraction of the user interfaces of the system, it has the responsibility to receive interactions from the user (and relay them to *presenter*) and display data sent to it from *presenter*. In our project, we implement Sleep

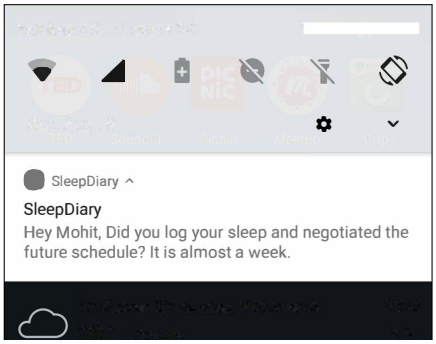


Fig. 4.10.: Model-View-Presenter Architecture Diagram

Entry UI, Summary UI, Negotiation Wizard UI, etc. All the user interactions are communicated to the *presenter* which handles the interaction.

- **Presenter**: This layer acts as the connecting layer for the model and view. All the user interactions are communicated to this layer which manages the querying the *model* and updating the *view* if and when needed. For our application, we needed a manager to handle the diaries, one to manage the summary and one to manage the recommendations.

The crucial part is that the presenter works as the middleman in both models to view and view to model communication, so it has the responsibility to interact with multiple models at times, but at a time a single view at a time.

Our existing application is built on the same architecture; it can be seen in by the component diagram in Figure 4.11 which indicates how components interact with each other.

In the Model layer, we had two components:
(1) Sleep Data
(2) Negotiation Data
Some presenters needed both the data, while some just needed one of those. For example, Diary Manager responsible for handling sleep diary just needed access to

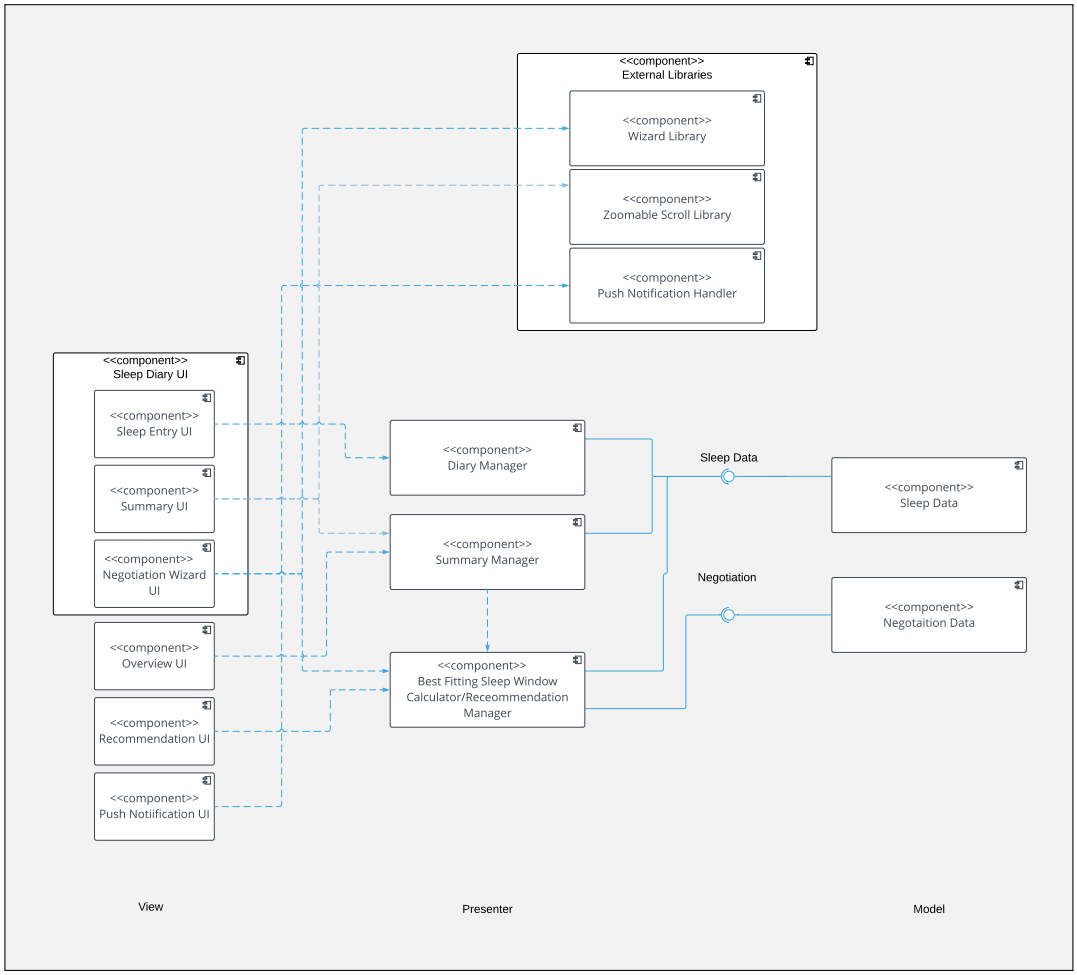


Fig. 4.11.: Component Diagram indicating the MVP architecture we followed to create the application

Sleep Data, while Best Fitting Sleep Window Calculator needed access to both types of data to predict the best fitting sleep window.

In the Presenter Layer, we had three components:
(1) Diary Manager: handling the adding and removal of diary entries.
(2) Summary Manager: handling the summaries generated from Sleep Data.
(3) Best Fitting Sleep Window Calculator/Recommendation Manager: handling the best fitting sleep window and manages the existing negotiated data.

In the View Layer, we had four significant components as the user interfaces(UI):
(1) Sleep Diary UI: This view has three subcomponents:
(1.1) Sleep Entry UI: The view to interact with sleep entry.
(1.2) Summary UI: The view to interact with the 7-day summary.
(1.3) Negotiation Wizard UI: The view to interact with, to negotiate sleep window.
(2) Overview UI: This view to interact with the complete summary of all sleep data captured in the app.
(3) Recommendation UI: The view to interact with to get the negotiated schedule.
(4) Push Notification UI: The view to interact with to push notifications received for the app.

We used a few libraries for implementing a few functionalities for faster prototyping.
(1) Wizard Library: The libraries for wizard-style interaction for negotiation of sleep schedule [Mar18].
(2) Zoomable Scroll Libary: The library used by Summary UI for carousel-like display for the sleep diary entries [Yar18].
(3) Push Notification Library: To handle the push notification, we used push notifications library for AppCenter [Mic18].

As described before the view layer never interacts directly with the model layer. All the view layers do is it delegates the presenter if the data is needed so that it stays open for further interactions, while the presenter layer fetches the data and make it handleable for the view layer. Using this architecture, another app can be created in other technologies as well and this architecture is not limited to Android.

4.4 Conclusion

This chapter explained how the mobile application Sleep Scheduler came into existence. After explaining the background, we moved on to discussing the design process we followed during the creation of the android prototype, after which we explained how we implemented the app, albeit on a high level. This chapter should serve as the starting point for similar apps if they are looking to develop or design such apps. In the upcoming chapter, we discuss the methodology of our research.

5 Research Methodology

> *Methodology should not be a fixed track to a fixed destination but a conversation about everything that could be made of happen.*
>
> **— J. C. Jones**
> (Design Methods, 1970-1992)

The purpose of this study was to understand the user experience of a sleep diary app called Sleep Scheduler and answer the questions we have raised in our research objective section (Section 1.3.2). To reiterate, Sleep Scheduler app is a digitalised version of the sleep diary, an instrument which in paper form is already in use by somnologists. The study which we did was done in two phases in which the users were involved directly, the ideation phase (focus groups) and validation phase (user tests). In Chapter 3, we talked about the focus groups at length. The following sections talk about the methods and techniques used during these user tests.

5.1 Participants

For the phase of the application evaluation, the user recruitment process was straightforward compared to the focus groups as we were limited by the number of phones (8) which we can test with. Thus, invitations were sent to the social circles who earlier expressed interest in being participants of the project. The selection criteria again were similar to the focus groups, the same demographic of 20-35 years old who spend over 10 hours per day on average in front of a screen. Different users were recruited in comparison to the focus group, to avoid bias which can arise due to prior knowledge of the mobile application.

There were eight males recruited based on a first-come-first-serve basis and their availability for the user tests. The sampling criteria were also similar to the focus group; they have to understand English in a professional context to understand the interview tasks, and can communicate their ideas. The participants were asked to sign an informed consent form, approved by ethics committee at human media interaction (HMI) department at the University of Twente and Philips. The project was also approved for privacy consideration by privacy committee at HMI department.

5.2 Research Pre-requisites

The user test phase of the study involved three parts:

1. Pre-interview,
2. Interaction with the app, and
3. Post-interview

The study in total(involving pre-usage interview, daily interactions, and post-usage interview) took 2 hours per participant. It was divided into pre- and post- interviews to evaluate changes before and after the application usage. The daily interaction amounted to 2 to 4 minutes per session on average in a day. The study was done for ten days.

The participants were informed of two things:

1. The app has no cloud component pulling their sleep data. It is always stored locally on the device and never pulled. We were interested in user experience and perception about the app.
2. There will be occasional push notifications sent to their phone, so for this, they needed to have an internet connection. They were provided with the phone for the duration of this project.

To record the quantitative data during the process, we used Credibility Expectancy Questionnaire (CEQ), Perceived Credibility Questionnaire (PCQ), Insomnia Severity Index (ISI) and few questions in the interview asking the participant to rank items (such as trustworthiness, expertise, and motivation) on a scale 1-10. To record the qualitative data, interview questions were asked and recorded. Later they were transcribed and analysed, and we discuss their analysis in the Chapter 6. We will discuss these scales and questionnaires in upcoming subsections.

5.2.1 Screening Scales

Insomnia Severity Index

Insomnia Severity Index (ISI) is a commonly-used screening mechanism for patients with insomnia. It is also used as the outcome measurement instrument of sleep therapy. The instrument has seven items assessing the quality of sleep. The one version we used was a self-reporting version. Two parallel versions exist, to be filled by the clinician and spouse respectively. The self-reporting version is put in Appendix F.

In [BVM01], Bastien et al. evaluated the psychometric properties of the tool and presented the assessment that the tool has high internal consistency and adequate concurrent validity.

The seven items ISI comprises are about the severity of sleep difficulties (onset, sleeping late, waking early), satisfaction with sleep pattern, perceived impairment in daily functioning, the perception of current sleep issues, and perception about sleep's role in the quality of life. The seven items were rated on a 0-4 scale, and the total ISI score varies from 0 to 28.

The guidelines from [BVM01] for scoring are:
Total score from 0-28:
0-7 = No clinically significant insomnia,
8-14 = Subthreshold insomnia,
15-21 = Moderate Clinical Insomnia, and
22-28 = Severe Clinical Insomnia.

Here are our 8 participants anonymised as P1..P8 are on the Insomnia Severity Index scale, which was taken before them starting the use of the app. Except for P6 who had moderate clinical insomnia by self-reporting on this index, every other participant has sub-threshold insomnia or no clinically significant insomnia Figure 5.1. P6 reported it might be due to their contemporary schedule modifications, and they felt healthy.

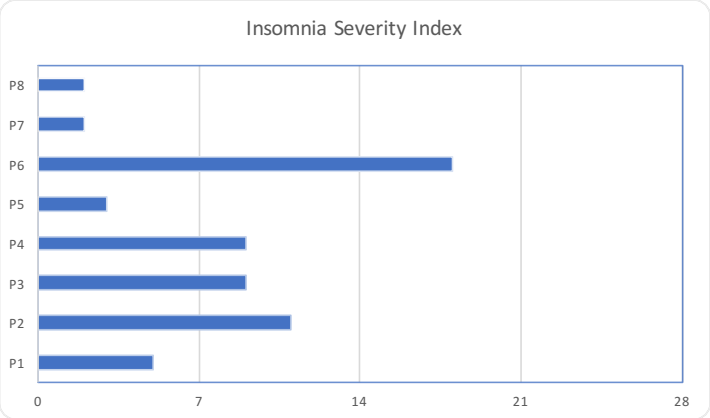


Fig. 5.1.: ISI criteria for the participants, P6 was noticably high on this scale

5.2.2 Credibility Evaluation Scales

Credibility Expectancy Questionnaire

Credibility Expectancy Questionnaire (CEQ) is a questionnaire used to evaluate the rationale credibility and outcome expectancy a clinical treatment [DB00]. The questionnaire consists of 2 set of questions. Set 1 has questions about what the participant in the treatment *thinks* and set 2 has questions about what the participant *feels*. Set 1 has four questions, and Set 2 has two questions. Thus there are six questions in total. The questionnaire is put in Appendix G.

In the study done to psychometrically evaluate the questionnaire by Devilly et al. [DB00], they did multiple studies. They presented this assessment of the factor structure for the CEQ: the first three questions map well to the credibility and the last three to the expectancy factor. A high standardised α between 0.84 and 0.85 was obtained for the scale (Cronbach's α of between 0.81 and 0.86 for the credibility factor and standardised α of between 0.79 and 0.90 for the expectancy factor). Thus, implying that the scale has high internal consistency and concurrent validity.

There are two rating scales in the CEQ, 1 to 9 for the questions 1, 2, and 3 of set 1 and question 1 of set 2 and 0-100% for the question 4 of set 1, and questions 2 of set

2. As the author of the questionnaire recommends to standardised the measure, we scale linearly the values 0-100% to 1-9 for evaluation.

Perceived Credibility Questionnaire

Smith and Vogt analysed negative word of mouth communication for ad processing, one of the dependent variables they observed was perceived credibility. They were using these questions for an advertisement (a non-computational product). We adapted a few questions to capture the mobile app's perceived credibility based on their presumption (presumed credibility).

PCQ-1 How truthful do you think the app is?
Not at all truthful (1) to completely truthful (7)
PCQ-2 How accurate do you think the app is?
Not at all accurate (1) to completely accurate (7)
PCQ-3 How credible do you think the app is?
Not at all credible (1) to completely credible (7)

During this project, we intend to see the influence of credibility on adherence. We recorded the perceived credibility questionnaire based on first impressions.

User Experience Questionnaire

User Experience Questionnaire (UEQ) is an instrument for recording user experience created by Laugwitz, Held, and Schrepp [LHS08]. This tool allows for simple and fast data collection on 26 items efficiently categorised on six scales. The same study validates the scales as well. The questionnaire can be found in Appendix H.

In another study by Schrepp et al. [SHT14], that measured the psychometric properties of this scale through 11 usability tests and found high reliability (measured using Cronbach's α).

Each of the items was to be given a score between 1-7. The item on the left indicates a 1 and on the right indicates a seven as the score. The six scales along with the corresponding items are as follows:

1. Attractiveness: calculates the overall impression of the product
   Items:
   annoying <> enjoyable
   good <> bad
   unlikeable <> pleasing
   unpleasant <> pleasant
   attactive <> unattractive
   friendly <> unfriendly
2. Perspicuity: calculates the learning required to use the product
   Items:
   not understandable <> understandable
   easy to learn <> difficult to learn

complicated <> easy
clear <> confusing

3. Efficiency: calculates how efficient is the product in the primary task of the user.
   Items:
   fast <> slow
   inefficient <> efficient
   impractical <> practical
   organized <> cluttered
4. Dependability: calculates the user perception of their control.
   Items:
   unpredictable <> predictable
   obstructive <> supportive
   secure <> not secure
   meets expectations <> does not meet expectations
5. Stimulation: calculates the stimuli introduced by the product.
   Items:
   valuable <> inferior
   boring <> exciting
   not interesting <> interesting
   motivating <> demotivating
6. Novelty: calculates the innovation and creativity as the user perceives.
   Items:
   creative <> dull
   inventive <> conventional
   usual <> leading edge
   conservative <> innovative

This scale can be used to compare multiple products on user experience, also it can be used to evaluate a before/after experience of a specific product (which is our case).

5.3 Process

The user tests were arranged to evaluate the user experience concerning *RQ 1* and *RQ 2* . *RQ 1* is about measuring the credibility of the app, while *RQ 2* is about the adherence. There are a few interview questions along with the instruments mentioned in the previous section which are used to get the answers to these questions. For the user-tests, we ran a pilot internally before going to each phase, which is a good practice to understand and possibly counter avoidable circumstances.

This section will indicate the process we followed for conducting the user tests. The process we followed is indicated in Figure 5.2. It is formed of three sub-phases: Pre-App Usage Assessment, App Usage, and Post-App Usage Assessment. The sub-tasks are indicated on the right of each of the phase indicated. Now we discuss individual phases of user tests.

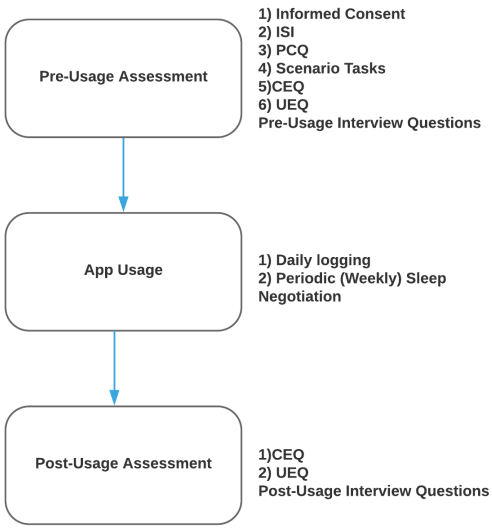


Fig. 5.2.: Process for User Tests. The sub-tasks are indicated on the right of each of the phases.

5.3.1 Pre-Usage Assessment

Informed Consent

The pre-usage assessment was in total a 30-40 min phase. In this phase, the user was welcomed to the user tests and was given an informed consent form. Once they have all their doubts about the study cleared, they were asked to sign the form. Once their informed consent was received, we shared the Insomnia Severity Index form.

Insomnia Severity Index

The participants were given the insomnia severity index scale after this consisting of 7 items regarding their sleep. This was a way to capture their current sleeping habits and if they faced any issue.

Perceived Credibility Questionnaire

Now we showed them the app on one phone; they were asked to look around in the app for about 30 seconds. Then the first questionnaire was given, this was the perceived credibility questionnaire (PCQ) to record their first impressions for the presumed credibility, this is to document what is the user's credibility due to presumption about the product.

Scenario Tasks

Once they completed the PCQ, they were given two tasks to do so that they are acquainted with two essential tasks- adding a sleep diary and negotiating a sleep schedule. The scenario tasks were as follows (they were instructed verbatim). If they were stuck at any part they were asked to explore the other buttons, their responses

were noted, and the audio was being recorded. The issues people reported were on the usability of the app the information was shared with the research team after the pre-usage assessment to plan for further improvements in the app.

Scenario Task 1 was done to acquaint the user with the task of adding a sleep diary. Thus they were given a fresh install of the app. Scenario Task 2 was done to familiarise the user with the task of sleep window negotiation. Sleep schedule negotiation is invoked once the user adds the 7^{th} entry, the negotiation interaction was discussed in detail in Section 4.3.1.

Scenario Task 1: Diary Addition
It is all about adding records and on-boarding to the app. Be open and think aloud while using the app, so that I can understand the thought process.

1. *Please open the app and try to add a new entry.*
2. *The diary entry should indicate this:*
   2.1. Sleep from 00:00 to 8:00.
   2.2. Nap from 14:15 to 15:15.
   2.3. Awake in bed from 11:30 to 12:00.
3. *Remove the nap time by using the awake and update it to 14:15 to 15:00.*
4. *Save the diary once you are done.*

Scenario Task 2: Schedule Negotiation

1. *We have entered six entries for you.*
2. *Add another entry similar to your last day.*
3. *If you see a pop-up appear, please press NOW to go to next step.*
4. *In this step,*
   4.1. Please see whether you find out the meaning of best fitting sleep window shift.
   4.2. Try to navigate around to see 7-days summary.
   4.3. Once you have found out, move on to next step.
5. *Now you see a new screen, set the sleep duration to 6:45 and press PROCEED to go on next step.*
6. *The app calculates a new preference based on your current choice. The red bar indicates your choice of sleep duration you have at the best fitting window, drag it around to see if you can adjust it to 11:45 to 6:30.*
   Once you are done, please press submit to move to next step.
7. *Let's say you did not like the timing you just chose. Now you can see the recommended sleep hours and the algorithm suggested in green; press PROPOSE AGAIN to go back and change the timing to 12:00 to 6:45.*
8. *You are still not happy with the recommended schedule, press PROPOSE AGAIN to go back and reset to 12:15 to 7:00.*
9. *You like the coverage now, and press accept to choose the recommended sleep window.*
10. *Now it shows you the final schedule, which is from 12:15 to 6:45. Press submit to conclude.*
11. *We still do not refresh the app, so please go out and come back in the app and see the updated assignment.*

Credibility Expectancy Questionnaire

Once the participant performed the scenario tasks, they were given the credibility expectancy questionnaire (CEQ) to evaluate our app based on the scenario tasks they just completed.

User Experience Questionnaire

After CEQ, the user experience questionnaire (UEQ) was filled by the participants to record the user experience based on the scenario tasks. Once they were done, they were asked the following interview questions.

Pre-Usage Assessment Interview Questions

1. Did you notice the Philips Logo on the launch screen. What do you feel on seeing that, and how would it change your perspective in comparison to something which doesn't have such a logo? (Reputed credibility)
2. Do you believe the styling, typography and visual appeal is suitable for the task of sleep logging by manual? (Surface credibility)?

The answers for each of these were recorded in paper form then digitalised, and the interview questions text were recorded and later transcribed.

The data was aggregated and later analysed. Each participant was given a phone, and the phone token for setting up the notification was noted so that we can set up the push notifications.

5.3.2 App Usage

For the participants, the interaction generally varies from 1-2 minutes a day just entering their schedule, and on the day they negotiate their sleep schedule, it goes up to 5 minutes.

In addition to these, we sent five push notifications. They were posted on Day 0 (Welcome), Day 1, Day 3, Day 4, and Day 7. The Day 0, 1, 4, and 7 were all weekdays. Day 3 was a Sunday. The notifications were sent in the evening between 4 p.m.-8 p.m. for the weekdays while around 1 p.m. on weekends. The personalisation in the messages is just in the form of the name of the participant being placed explicitly in the placeholder **<Participant Name>** in the messages.

The five messages were as follows:
Day 0: Welcome to our user test. :)
Day 1: Hi *<Participant Name>*, Hope you had a nice sleep, fill those hours in the sleep scheduler app soon if you haven't already.
Day 3: Hey *<Participant Name>*, Did you sleep enough Saturday night? National Sleep Foundation recommends 7 to 9 hours of sleep for your age group. Don't forget to make an entry in the sleep scheduler.
Day 4: A sleep tip, just for you, *<Participant Name>*. Your body needs time to shift into sleep mode, so spend the last hour before bed doing a calming activity such as

reading. Source: National Sleep Foundation [1]
Day 7: Hi <*Participant Name*>, Did you log your sleep and negotiated the future schedule? It is almost a week.

During this phase, the users only had to log their sleep data on the Sleep Scheduler app and no questionnaire or survey was to be filled.

5.3.3 Post-Usage Assessment

Similar to the pre-usage assessment, the post-usage assessment was also around 30-40 minutes long.

Credibility Expectancy Questionnaire

As now the participants were quite aware of the app use. They were welcomed and discussed generally the app; they were given the credibility expectancy questionnaire(CEQ). This was done to evaluate post-usage assessment credibility of the application.

User Experience Questionnaire

The CEQ was followed by User Experience Questionnaire(UEQ). This was used to record the user experience after using the app for a while so that we can understand how the user experience changed over time. The interview questions followed this.

Post-Usage Assessment Interview Questions

The interview questions were a mix of qualitative and quantitative ones as the users were asked to rate the app on various scales and the reasons were asked for giving the rating.

1. What is your general impression of the app
2. On a scale of 1-10, one meaning not at all motivated and ten meaning highly motivated, how motivated were you to use the app?
   a. Can you tell me why you gave that rating?
3. Can you name the reasons that made you keep using the app?
4. On a scale of 1-10, one meaning not at all trustworthy, and ten meaning trustworthy, how would you rate the app?
   a. Why did you choose this?
   b. Has it contributed to the motivation to use the app?
   c. If so, in which ways
5. On a scale of 1-10, one meaning not at all expert, and ten meaning an expert, how do you perceive the expertise of the app based on the advice it provides for sleeping on a specific schedule?
   a. Why did you choose this?
   b. Has it contributed to the motivation to use the app?
   c. If so, in which ways

[1] https://sleepfoundation.org/press-release/national-sleep-foundation-recommends-new-sleep-times/page/0/1

6. On a scale of 1-10, one meaning not at all relevant, and ten meaning completely relevant, how relevant did you find the personal notifications?
   a. What role did the personal play in your decision to log?
7. Were there any other factors that influence your motivation to log?
8. What are your general impressions of the negotiation feature?
   a. There were a few personalised notifications for your progress, sleep facts, and some tips about good sleep. Which one out of these would help you adjusted to a new sleep schedule as recommended? Why?
9. Here are the few notifications which I sent. They consist of your progress, sleep tips, and sleep facts. Can you put these 3 in order in which they influence you to use the app?

   [Progress]: Hi <Participant Name>, Hope you had a nice sleep, fill those hours in the sleep scheduler app soon if you haven't already.

   [Sleep Fact]: Did you sleep enough Saturday night, <Participant Name>? National Sleep Foundation recommends 7 to 9 hours of sleep for your age group. Don't forget to make an entry in the sleep scheduler. Source: National Sleep Foundation

   [Sleep Tip]: A sleep tip, just for you, <Participant Name>. Your body needs time to shift into sleep mode, so spend the last hour before bed doing a calming activity such as reading. Source: National Sleep Foundation [2]

The answers to these interview questions were recorded. Later they were transcribed, coded, and then analysed using thematic analysis (as described in Section 3.3.1). Unlike the focus groups (where the design features were the themes), we did not have a set of themes to examine. The thematic analysis followed the guidelines from Braun and Clarke involving discovery of themes [BC06]. To validate this, two researchers on this project explore the themes individually, and by consensus, reached the final themes and sub-themes. The coding was done using the tool Nvivo similar to focus groups [Ltd18].

5.4 Conclusion

In this chapter we discussed the methodology we followed during the project, specifically during the user tests. We started by introducing how the participants were recruited, the instruments we used to measure or obtain the data, and the procedure we followed to conduct our user tests. The coming we will discuss the results obtained during this project.

[2] https://sleepfoundation.org/press-release/national-sleep-foundation-recommends-new-sleep-times/page/0/1

Results

6

> *As a scientist, you're not supposed to make decisions without the data.*
>
> **— Francis Collins**
> American Scientist, Director at NIH

The chapters before this discussed in depth how sleep scheduler was set up as a tool for research; this chapter delineates the results obtained by the study. We present the various evaluation based on types of credibility (presumed, reputed, surface, and earned), adherence evaluation, and user experience evaluation.

6.1 Presumed Credibility Evaluation

As defined in Section 2.4, presumed credibility is the credibility a user perceives of a product due to the presumptions about it. We tried to evaluate this using the perceived credibility questionnaire which contains three questions about the following items: truthfulness (PCQ-1), accuracy (PCQ-2), and credibility (PCQ-3). The items were evaluated on scales of 1-7. For more details, please refer to Section 5.2.2.

We obtained high credibility scores for presumed credibility on all the three items as the participants rated the application to be at average scores for truthfulness at 5.375 (SD =0.916), accuracy at 5.125 (SD =1.126), and credibility at 5.625 (SD = 0.916) (see Figure 6.1).

6.2 Reputed and Surface Credibility Evaluation

As discussed in Section 2.4, reputed credibility is the credibility you gain due to the reputation of a third party endorsement. We used the Philips logo on the launch screen to indicate that this is a Philips project. Also, after the scenario tasks, participants had a first impression of the app, so we could evaluate surface credibility based on the looks and feel of the app.

Based on their initial interview we recorded their perception of reputed credibility and surface credibility. These were the two questions we asked in the pre-usage interview:

1. Did you notice the Philips Logo on the launch screen. What do you feel on seeing that, and how would it change your perspective in comparison to something which doesn't have such a logo? (Reputed Credibility)

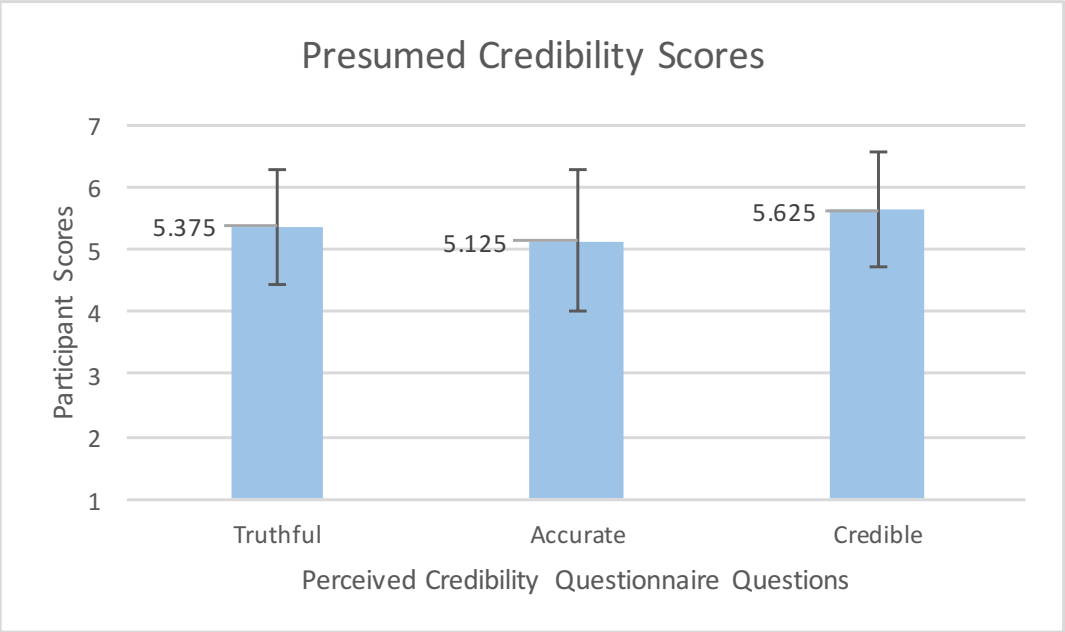


Fig. 6.1.: Perceived Surface Credibility for the participants

2. Do you believe the styling, typography and visual appeal is suitable the task of sleep logging by manual? (Surface Credibility)

Over half the participants noticed the Philips logo. And we found out the following keywords from the transcriptions as described in Table 6.1. The reputed credibility in such an app was synonymous with the brand of Philips. Few people found the colour blue hospital-like while others found it serene. Although all of them saw the app as unambiguous and user-friendly, there was a continuous reporting of fat finger error (EXPLAIN THIS TERM). A few participants expected advanced features such as night mode, help menus etc. as well.

Tab. 6.1.: Keywords from pre-interview for reputed and surface credibility

Part Id	Reputed Credibility	Surface Credibility
P1	Brand, Trust, First Impression, Inherent Credibility	User-friendly, Coloring, Easy to Understand, Not complex
P2	Past Experience, Well-designed	Unambiguous, Hospital-like
P3	Credibility, Recognizable	Fat Finger Error
P4	Minimal, Brand	Fitbit-like, Representative colors, Lack of few Helps/ Hints
P5	Brand	Blue color (calm/peace) , Fat Finger Error
P6	Enhances Credibility, Developer Information Needed	Daytime-friendly, No Night-mode
P7	Trust, Brand	Blue Color (serene)
P8	Brand, Well-tested	Help menu missing, Hints Old-fashioned

6.3 Earned Credibility Evaluation

Earned credibility is the credibility a tool gains after a continuous period of the first-hand usage. To capture this, we used the Credibility Expectancy Questionnaire. This is a questionnaire which is used to evaluate the credibility and expectancy of a medical treatment.

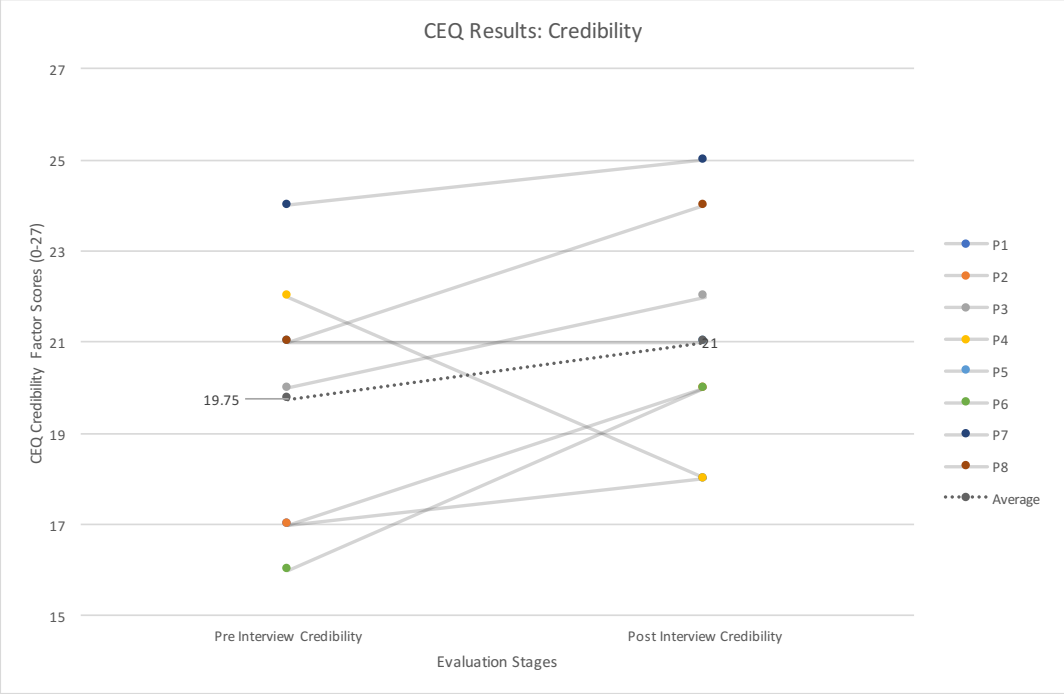


Fig. 6.2.: Comparison between credibility scores prior to and post usage

As there are two scales - 1 to 9 and 0 to 100 %, we scaled the 0 to 100% to 1 to 9 by linearly scaling. We present comparison of credibility between pre-app usage and post-app usage, although for P4 the credibility decreased over this period, on an average, upon usage, our application gained the (average) credibility and moved from 19.75 ($SD = 2.816$) to 21 ($SD = 2.563$) out of a maximum score of 27 (see Figure 6.2). We found that the (mean) expectancy of sleep scheduler as a medical tool reduced from 16.71 ($SD = 4.436$) to 15.585 ($SD = 3.880$) out of a maximum score of 27 (see Figure 6.3).

6.4 Adherence Evaluation

All 8 participants logged at least seven days out of 10 as were evident from the recorded sleep diaries, which was a considerable success. To figure out the reasons for adherence we had to analyse the post-usage interview questions. We will present the analysis next.

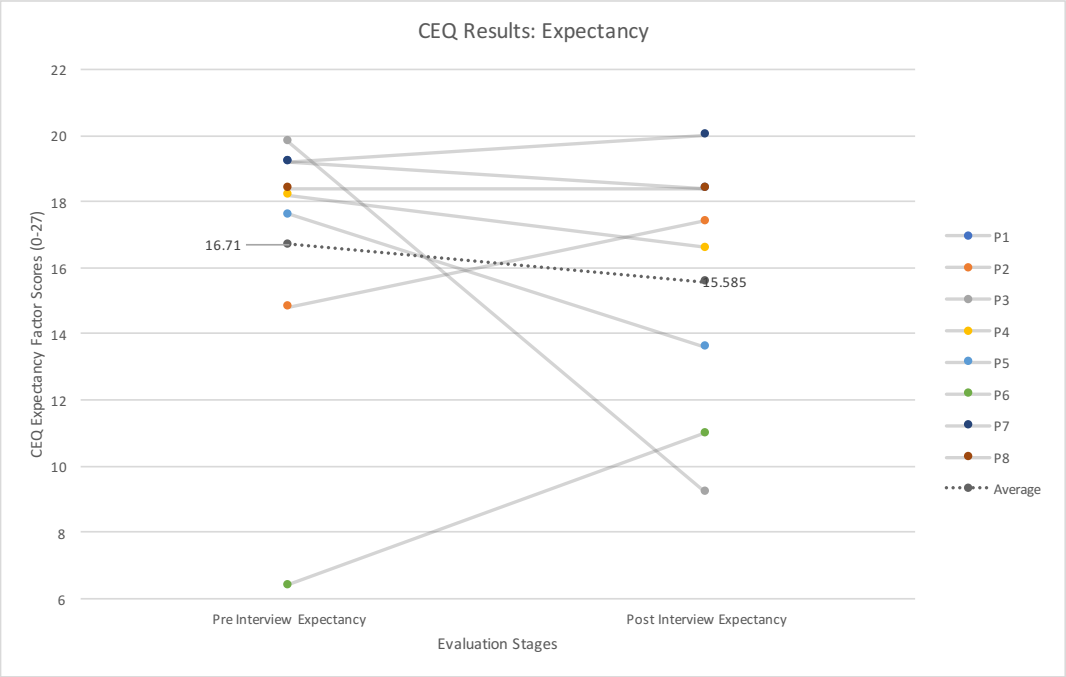


Fig. 6.3.: Comparison between expectancy scores prior to and post usage

6.4.1 Post-Usage Interview Questions Analysis

Quantitative Research

The final interview evaluates credibility as made of components - trustworthiness and expertise. Thus the scores for questions such as

1. On a scale of 1-10, one meaning not at all trustworthy, and ten means trustworthy, how would you rate the app?
2. On a scale of 1-10, one meaning not at all expert, and ten means an expert, how do you perceive the expertise of the app based on the advice it provides for sleeping on a specific schedule?

Additionally, we wanted to know more about their motivation to use the app. Thus we had this question in the questionnaire.

3. On a scale of 1-10, one meaning not at all motivated and ten means highly motivated, how motivated were you to use the app?

We did not get the value for expertise on 1-10 scale from P2, P6 so those values were ignored from mean and standard deviation calculation for the expertise. The results were are as follows: trustworthiness was scored at 8.625 ($SD = 1.060$), expertise at 7.167 ($SD = 0.983$), and motivation at 7.5 ($SD = 0.534$), indicated in (Figure 6.4). The scores were again indicative of good credibility evaluation for our app, especially in terms of trustworthiness.

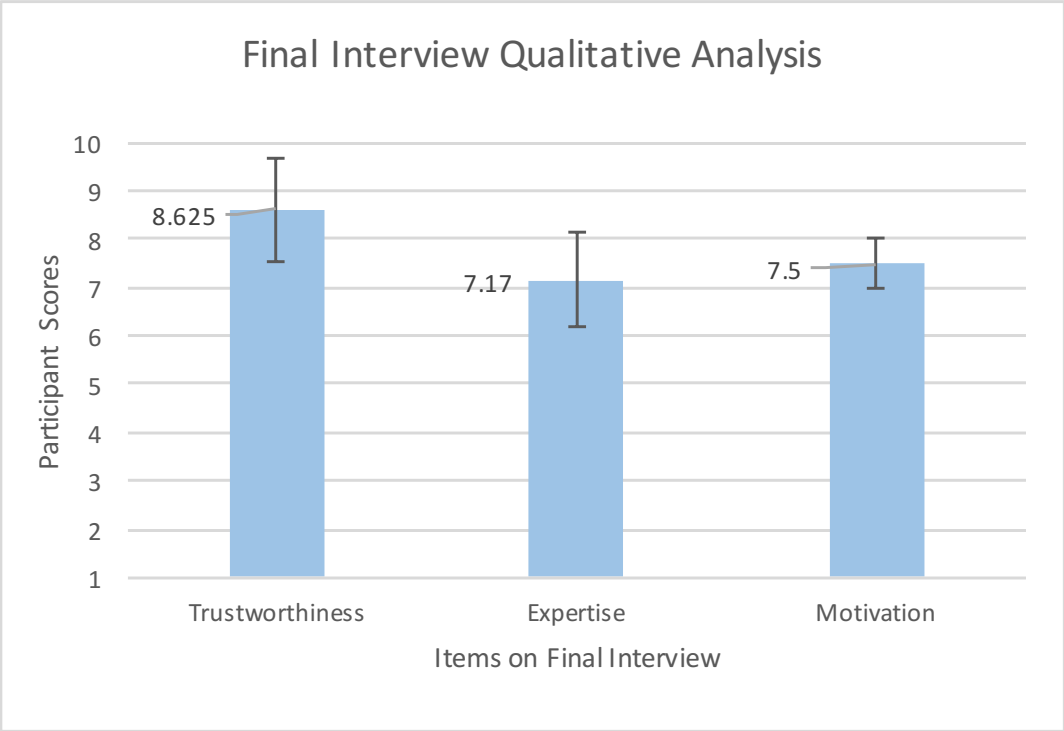


Fig. 6.4.: Mean and standard deviation for the quantitative data of final interview questions

Qualitative Research

The final interview questions were analysed qualitatively using thematic analysis, and the results are reported, similar to Section 3.3. The thematic analysis is given in Table 6.2. The main themes were:

1. User Experience
   Sub-Themes: Future Add-ons, Advantages, Usage, Difficulties
2. Trust
   Sub-Themes: Transparency, Look and feel, Expertise
3. Motivation
   Sub-Themes: Interest, Encouragement
4. Information.
   Sub-Themes: Relevancy, Personalization, Feedback

The three most prominent sub-themes were Interest, Future Add-ons, and Relevancy. The interest was expecting a result after putting some effort into learning how to use the app to yield the best returns. Also, there were lots of points of improvements discussed regarding feedback and notification on the future add-ons sub-theme, and the relevancy was indicative of the app is relevant to the roles it was created for, that is for logging sleep, and making thinking about sleep more cognitively.

Tab. 6.2.: Thematic Analysis of Final User Interviews

Subtheme (Coverage)	**Description**	**Example Quotes**
Theme: **User Experience**		

Continued on next page

Tab.6.2 – continued from previous page

Subtheme (Coverage)	Description	Example Quotes
Future add-ons (12.10%)	(1) Adding more feedback as: [a] comparing previous day's schedule [b] a summary of sleep routine [c] instant after filling schedule ("good", "well-done") [d] having the schedule more open for exceptions (e.g. weekends) [e] info about personal progress (2) Adding different kind of notifications [a] reminding about bedtime [b] informing about the importance of logging	*You can motivate me then when you put instant feedback of "well-done! You should have slept more," or as soon as I put I slept from four in the morning to six in the morning, it would be motivating if I would get a message saying "your sleep sucked tonight you should have you only slept two hours.".*
Advantages (10.14%)	(1) The app is sufficiently flexible for adjusting sleep time. (2) Negotiation is a playful interaction. (3) The app is easy to use, efficient and organised with a clear colour coding on states. (4) The advice on sleep window shift and push notifications are useful in that it adds a clear goal to the participant. (5) A digital sleep diary is more advantageous than physical, and it motivates to push notifications and advice.	*"The interaction was quite user-friendly and used to make sense that I cannot change so much from my schedule so suddenly."* *"I think it was good, it allows you to negotiate twice, and I did not try that much, I did try twice. I liked a schedule which I liked; I just switched it a bit backward."*
Usage (9.06%)	Push notifications help to remind filling; the app is usable during the daily schedule. The text amount in the notifications needs to be reduced. It takes a while to make it part of the daily habit.	*I use the sleep schedule app normally, tried to incorporate it into my daily routine so that I make sure I fill it and sleep on time. The time of day was not that important, but I just wanted to fill it so that I can be done with it.* *It became a kind of habit that I used to open the app and check so that I logged whenever I used to open the app.*
Difficulties (4.67%)	The first days perceived as confusing then it became easier. Using the phone on a second device. Having a big finger.	*If the notifications came to my primary phone, I would be better served by the purpose by motivating me to log.*
Theme: **Trust**		
Data Transparency (4.90%)	(1) The knowledge of data being kept secure. (2) Having Philips logo at the beginning of the app. (3) Not having ads or a third party involvement. (4) Being aware of user's boundaries and not overstepping to the data.	*I think that it was good and I knew that my data is not going to be shared with any other third party and being the logo of Philips was also making me also giving me a lot of different forms of trust, so I think perfect ten.*

Continued on next page

Tab.6.2 – continued from previous page

Subtheme (Coverage)	Description	Example Quotes
Look and feel (4.78%)	(1) Design and user-friendliness make it look credible. (2) Simple and has a clear goal.	*The looks made me use the app more, and it felt trustful, it looks worthy, and there wasn't like a lot of work to do. It is a simple purposefully built app; I don't like big apps. It is cool because it's for a specific job, it is trusted for that job.*
Expertise (4.60%)	(1) Source of the app is important. (2) The concepts being developed with knowledge of insomnia is important. (3) The app will be perceived as more credible if it is verified by a sleep medicine expert.	*I think the sleep is important, and I have the sensation the feel to sleep not the correct hours, and I would need someone who understands my sleep better and help me with that. The credibility is what the word which will be given by the app if a doctor recommends the app, I will trust this app 100%.*
Theme: **Motivation**		
Interest (14.07%)	(1) Receiving personalised information to feel motivated. (2) Filling sleep diary itself to get more consciousness about sleep. (3) Receiving a relevant prediction from the app. (4) Interest would increase if there were a medical reason to use the app. (5) Receiving prediction of sleep records and personal notification increased curiosity for the app.	*When you started receiving personalised stuff personal text about yourself, so you feel more motivated. .* *The person who wants to sleep on time will do it. If they are the tips are more relevant to them.* *I knew that I would get a prediction from the app (from the demo task) and wanted to get in a better schedule of sleep personally.*
Encouragement (4.99%)	(1) Receiving personalised notifications. (2) Being aware of one's own sleep situation. (3) Data transparency.	*I would have returned it quite low if the push notification was not there because people would still forget but because of the push notifications.*
Theme: **Information**		
Relevancy (10.52%)	(1) The prediction was a good match of the existing schedule (2) The predicted sleep schedule is somewhat shorter and later than predicted expectation. (3) Receiving notifications.	*I think the prediction was quite accurate, but then I negotiated a little bit because it was a bit late because during this week I was sleeping a bit late.* *Final negotiation was a decent approximation of when I wanted to sleep, and the app recommended me to sleep.*
Personalization (4.86%)	(1) Personalizing with name is preferable. (2) Receiving personal progress is valuable. (3) Informing about the schedule. (4) Personal goal indication.	*It will be easier to sleep when you have taken the choice out of my mind of thinking when to sleep. I think the personal goal indication is the most important point to increase the adherence and program in general.*

Continued on next page

Tab.6.2 – continued from previous page

Subtheme (Coverage)	Description	Example Quotes
Feedback (4.47%)	(1) Tips helped to teach info. (2) The advice was sometimes different from the expectations. (3) Sleep window shift value was perceived as advice. (4) Adjusting sleep on negotiation screen helped to sleep on the recommended hours.	*And I was happy with the recommended schedule although it was not the real match of what I wanted as my schedule for that was a bit shifted, so I moved the sleep on my own. Something about adjusting the sleep was on the negotiation screen, and that one has helped me the most to try to sleep to the recommending sleep hours. Also, the tip about getting prepared for sleep.*

6.5 User Experience Evaluation

We also evaluate the user experience using the user experience questionnaire. The 1-7 scale of the user experience questionnaire (UEQ) was scaled to -3 to 3 for analysis. A tool available from the authors did the analysis[1]. The results are indicated in Figure 6.5.

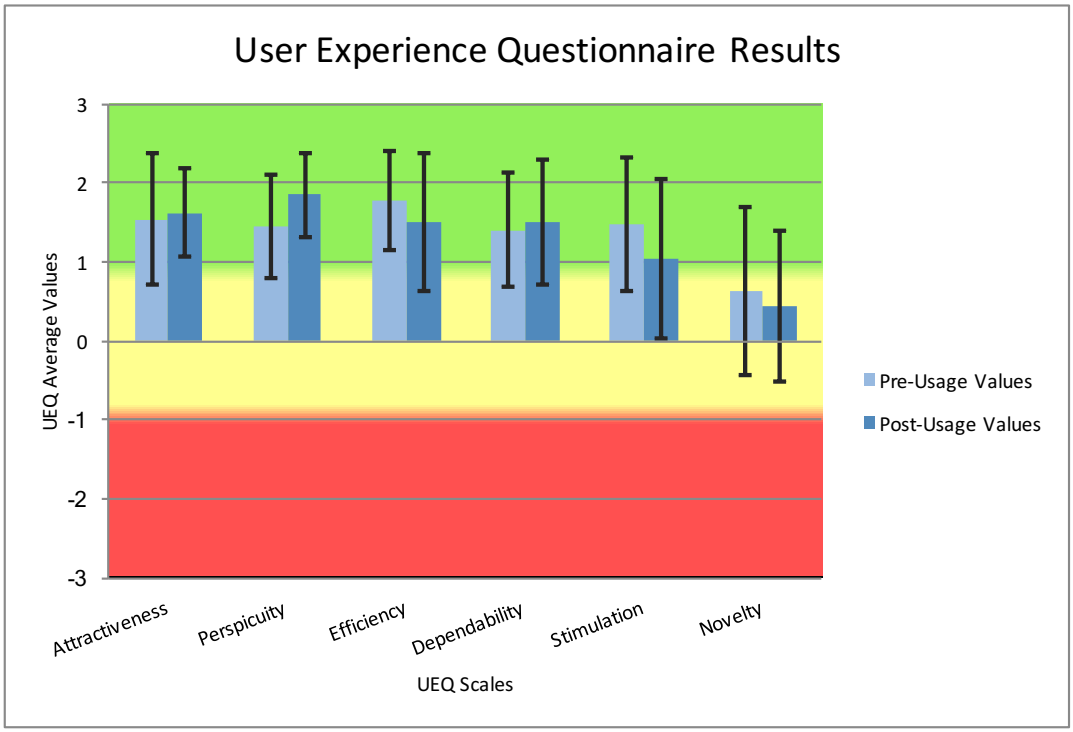


Fig. 6.5.: Comparison between user experience scores prior to and post usage. Green, Yellow, and Red colors indicate where whether the scale had positive evaluation, netural evaluation, and negative evaluation respectively

The 26 items of UEQ were combined to 6 scales, and the values before and after the usage are calculated. UEQ does not calculate a complete usability score since the

[1]The tool can be accessed at https://www.ueq-online.org/Download/FileDownload?id=3

interpretation of such an amount is difficult. The scales vary from -3 (horribly wrong) to +3 (extremely good). The values between -0.8 to +0.8 indicate neutral evaluation, greater than +0.8 indicate positive evaluation, and below -0.8 indicate negative evaluation. On comparing the pre and post usage experience evaluation, we found that they found the tool more attractive, more perspicuous, and more dependable than before, although the efficiency, stimulation and novelty were few scales where our tool performs poorly.

6.6 Conclusion

This chapter presented the analysed data from the user tests we conducted for sleep scheduler. We evaluated credibility as its different types - presumed, reputed, surface, and earned credibility. Then the analysis of the post-usage interview questions was discussed. The user experience evaluation followed this and found that we did well on some user experience scales while not in others. Next chapter discusses these results in the context of our research questions.

7 Discussion

> *A scientist's aim in a discussion with his colleagues is not to persuade, but to clarify.*
>
> — **Leo Szilard**
> American Scientist

The findings of this research studies suggest that the Sleep Scheduler app does help with the two tasks it was set out to do. Firstly, expediting sleep diary logging and secondly, provide the stimuli to think about sleep more cognitively, although to a moderate degree. The previous chapter indicated a positive inclination on the credibility scales, the importance of branding, and outlined the appreciation for colour choices for the app. CEQ also indicated a positive credibility change after the app was used for a while. Now we discuss the result in the context of our application.

7.1 Research Questions: Answered

Through this research, we wanted to figure out the most prominent persuasive design feature in the PSD Model [OH09] which can influence the credibility of an app. The goal of this research was to figure out the credibility levels of Sleep Scheduler (RQ 1) and the role of the credibility in adherence (RQ 2). We set out to get the answers to the research questions asked in Section 1.3.2.

DQ 1 What persuasive design features from the PSD model (Figure 1.4) can enhance the credibility of the sleep scheduler app?

The exploration into the PSD model was primarily a part of the ideation phase of this project. During this phase, we conducted three focus groups and from the qualitative analysis of the discussion text and the participant's proclivity to one feature in specific, *Personalization*. The details about the focus group and its data analysis are presented in Section 3.2.
The algorithm to decide the sleep schedule provides personalisation due to its adaptive nature to the individuals sleeping hours. To explicitly personalisation in the app, we used only the name of the participant in the notifications.
More than half of the participants noticed their name in the notifications and felt it's role to be important. To bring it in the implementation without altering the functionality of the game much, we created personalised notifications.

RQ 1 How credible do the users find the app to facilitate managing their sleep window?

This subquestions were:

SQ 1.1 Which credibility scale to use to evaluate the user's choice?

Although credibility has no single standard scale to capture in its entirety. We tried to achieve this in two ways. Firstly, by evaluating individual types of credibility such as presumed credibility, reputed credibility, surface credibility, and earned credibility. Secondly, by assessing the two dimensions which make up credibility: trustworthiness and expertise (see Figure 2.8).

The presumed credibility was recorded using the perceived credibility questionnaire (modified for mobile apps). We asked interview questions for the reputed and surface credibility and used credibility expectancy questionnaire (CEQ) for evaluating the earned credibility. More details can be found from Sections 6.1 to 6.3.

To evaluate trustworthiness and expertise quantitatively, we used our post-usage interview questions by asking participants to rate these items individually on a scale of 1-10. More details can be found in Section 6.4.1.

SQ 1.2 What is the user's evaluation of the app for credibility on the chosen credibility scales?

The evaluation of the sleep scheduler was really good in terms of the presumed credibility. On a scale of 1-7, the individual items on trustworthiness (PCQ-1), accuracy (PCQ-2), and credibility (PCQ-3) were scored 5.375 ($SD = 0.916$), 5.125 ($SD = 1.126$), and 5.625 ($SD =$0.916) respectively. The application follows material design guidelines [Goo18]; thus it would have looked similar to the apps participants have interacted with.

The reputed credibility was gained by belonging to a Philips project and was indicated by the Philips logo which was noticed by more than half of the participants. The feedback from the participants was unanimous in considering the brand presence of great importance and indicative of a well-tested product.

The surface credibility has mixed evaluation. Most people liked the colour choices for the app, but one of them find it hospital-like. Many of the participants were concerned about the error in logging due to the fat finger error.

To evaluate the earned credibility, CEQ was used, the app was scored (on average) at 19.75 ($SD = 2.816$) out of 27 before the app usage for ten days, and 21 ($SD = 2.563$) out of 27 after the app usage. This indicates that the app was perceived as more credible as there was continuous usage.

On the evaluation of credibility expressed as its dimensions of trustworthiness and expertise, the scores were again high. Trustworthiness was evaluated as 8.625 ($SD = 1.061$) out of 10 while on the expertise the rating was rated as 7.17 ($SD = 0.983$) out of 10.

The implication is that Sleep Scheduler had high credibility and was in general found credible with the information recording and processing.

Based on whether we evaluate credibility as its types or we evaluate it as made of its dimensions, the app scored high even though we can not express credibility as a single value from these results.

RQ 2 As we facilitate the user in logging their sleep and wake hours. How much is the adherence to this task?
The sub-questions were:

SQ 2.1 Is there a continued adherence to the task of logging?
Yes, there was a continued adherence with all of the participants logging for over seven days out of 10. To understand what made the participants log every day, there were few reasons outlined; this was obtained as a result of the thematic analysis of the interview.

- **User Experience**: The daily effort to enter diaries was minimal while the weekly result of getting a personalised sleeping schedule seemed a good enough motivator for our participants. The push notifications played the role of reminder and agent of personalisation (for the participants who noticed their names).
- **Trust**: The knowledge of this being Philips sponsored project, the information that their sleep data is not leaving the app and the no advertisements on the app created trust. This made them use the app regularly. While the simple, user-friendly design provided credibility, the participants also outlined that the credibility of the app will be more if a sleep therapist recommended it to them.
- **Motivation**: The increased cognition about their sleep, the personalised prediction of schedule, plus receiving personalised notification gave the participants strong motivation to keep using the app.
- **Information**: The relevance of the personal notification, smart prediction of personalised schedule, valuable sleep progress tracking, and a personal goal indication (negotiated schedule) were strong information parameters which made the participants use the app regularly.

There was feedback in the form of the recommended/negotiated schedule and the notifications, but participants wanted more funny feedback and wanted to use it more as an alarm clock/todo list before sleep, which they believed would have increased their usage of the app.

SQ 2.2 What was the role of perceived credibility in adherence to the task?
Perceived credibility was one of the critical parameters indicated by the themes *Trust* in the thematic analysis (Section 6.4.1). The key elements were the data transparency (no ads or data fetch out of the device), look and feel (Philips logo, user-friendly design and colour), and the expertise (in personalised sleep window prediction) which our application showed.

The role of credibility is important in a technology getting accepted, as told by Fogg in chapter 6 of the book Persuasive Technology ([Fog03]). Its role in adherence was monitored in this study, and we found its role unmissable. It can thus be safely implied that for a mHealth app like ours, credibility is one of the critical parameters.

7.2 Implication and Limitations

The research using the mobile application Sleep Scheduler was centred around the credibility evaluation of mobile apps. The literature research indicated the credibility

has an important role in behaviour change [Fog03]. The outcome of this study has just emphasised the same.

The Fogg's Behavioral Model can be another lens to look at the result. The equation

$$B = m \times a \times t$$

where B = target behaviour; m = motivation, a = ability, and t = trigger (or cue) for target behaviour. We brought the behaviour of changing the task to be so small as the ability to do this tiny habit (taking the phone out and putting the sleep hours) was achieved. The motivation was also high (as the participants were all volunteers). The anticipation of their sleep schedule provided the triggers, plus the personalised push notifications, so we ensured all three variables are present for continued adherence and were successful in our goals.

The limitation is that we did achieve a new habit being done for a fixed period (a green span behaviour according to the Fogg's Behaviour Grid, see Section 2.3.1), but it did not guarantee further continuous usage (blue path behaviour according to Fogg's Behaviour Grid, see Section 2.3.1). It is a real challenge which we could not solve during this study.

For the focus groups, we had six males and six females in the total 12 participants while for the final user test, all participants were males. This might have introduced a bias due to gender. Additionally, a few of the participants were from the primary researcher's social circle, and that might have added a bias. In future attempts to recreate the same experiment, these biases should be overcome.

7.3 Future Work

Future research should further develop and confirm these initial findings by running the user evaluation studies on a larger scale, and over a more extended period (through randomised-controlled trials), as then we will be able to provide statistically significant inferences.

Broadly translated our findings indicate that the presence of personalised information as a summary in the app and the personalisation in the notification both helped with the adherence. Future studies can explore what is the role of each of these kinds of personalisation individually.

The push notifications we used were quite elementary and limited by the names of participants. In the future, they can be evaluated concerning content, capability, and usability. The role of push notifications in mHealth apps, in general, should also be an interesting further area of research.

Conclusion

8

> *A conclusion is a place where you get tired of thinking.*
>
> — **Arthur Bloch**
> American Author

In this report, we have described the research study for Sleep Scheduler, an app which facilitates the log the sleep hours of people who have insomnia. It is a digitised version of a paper sleep diary (Section 1.2.1). Our goal was to include persuasive design features while creating this application and later evaluate the role of credibility in daily logging.

The end user(s) were involved in two phases - during ideation in the form of focus groups (Section 3.2) and during app evaluation. With the focus groups, we answered *DQ 1*, as we figured out that Personalization is the critical feature that persuades user a great deal in case of a mHealth app targeted to facilitate good sleeping habits (how we reached this conclusion can be found at Section 3.3). These learnings were implemented in Sleep Scheduler by personalising the push notification sent to the user while they were testing the app. The app was personalised by choice of algorithm, which customised the recommended schedule based on the user's sleeping schedule.

The *RQ 1* is answered as the participant found the app highly credible, various scales rating the app high on credibility, whether we evaluate credibility by types, or by its dimensions. The *RQ 2* was about the adherence, we found that all of the participants were quite adherent as they logged for over seven days out of total 10, we found that the role of credibility is critical as well. The discussion can be found in detail in Section 7.1. Along with these answers, the research provided a novel approach to compare credibility in general in mHealth apps.

This study sums up the credibility aspect for a mobile phone app called Sleep Scheduler, with this report answering a few fundamental research questions about credibility in the context of mHealth apps. The findings are quite useful for future apps in the mHealth industry to be checked for credibility.

Bibliography

[AAL13] Julian K. Ayeh, Norman Au, and Rob Law. “"Do We Believe in TripAdvisor?" Examining Credibility Perceptions and Online Travelers’ Attitude toward Using User-Generated Content”. In: *Journal of Travel Research* 52.4 (2013), pp. 437–452. DOI: 10.1177/0047287512475217. arXiv: arXiv:1011.1669v3 (cit. on p. 23).

[AS16] Alyssa Appelman and S Shyam Sundar. “Measuring Message Credibility: Construction and Validation of an Exclusive Scale”. In: *Journalism & Mass Communication Quarterly* 93.1 (2016), pp. 59–79. DOI: 10.1177/1077699015606057 (cit. on p. 23).

[BC06] Virginia Braun and Victoria Clarke. “Using thematic analysis in psychology”. In: *Qualitative Research in Psychology* (2006). DOI: 10.1191/1478088706qp063oa. arXiv: arXiv:1011.1669v3 (cit. on pp. 29, 54).

[Beu+16] Robbert Jan Beun, Willem-Paul Brinkman, Siska Fitrianie, et al. “Improving adherence in automated e-coaching”. In: *International Conference on Persuasive Technology*. Springer. 2016, pp. 276–287 (cit. on p. 12).

[BVM01] Célyne H. Bastien, Annie Vallières, and Charles M Morin. “Validation of the insomnia severity index as an outcome measure for insomnia research”. In: *Sleep Medicine* 2.4 (2001), pp. 297–307. DOI: 10.1016/S1389-9457(00)00065-4 (cit. on pp. 46, 47, 83).

[Cha+] J Chandrananth, A Fraval, YM Chong, M Chandranath, and P Tran. “A Guide to Writing eHealth Websites with Orthopaedics as an Example - Readability, Reliability, Usability”. In: *Journal of Orthopedics and Rheumatism* 1 (2), pp. 28–32 (cit. on p. 23).

[CS04] Deborah Charnock and Sasha Shepperd. *Learning to DISCERN online: Applying an appraisal tool to health websites in a workshop setting*. May 2004. DOI: 10.1093/her/cyg046 (cit. on p. 23).

[Dal+09] Meagan Daley, Charles M Morin, Mélanie LeBlanc, Jean-Pierre Grégoire, and Josée Savard. “The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers.” In: *Sleep* 32.1 (Jan. 2009), pp. 55–64 (cit. on p. 2).

[DB00] Grant J. Devilly and Thomas D. Borkovec. “Psychometric properties of the credibility/expectancy questionnaire”. In: *Journal of Behavior Therapy and Experimental Psychiatry* 31.2 (June 2000), pp. 73–86. DOI: 10.1016/S0005-7916(00)00012-4 (cit. on pp. 23, 47, 84).

[Dij05] Arie Dijkstra. “Working mechanisms of computer-tailored health education: Evidence from smoking cessation”. In: *Health Education Research* 20.5 (Oct. 2005), pp. 527–539. DOI: 10.1093/her/cyh014 (cit. on pp. 31, 34).

[Ert+18] Begum Erten-Uyumaz, Sebastiaan Overeem, Loe Feijs, Mili Docampo Rama, and Jun Hu. “A Digital Sleep Restriction System for Insomnia Therapy Based on Sleep Window Shift Negotiation”. In: *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference on - PETRA ’18*. New York, New York, USA: ACM Press, 2018, pp. 58–61. DOI: 10.1145/3197768.3201531 (cit. on pp. 3, 32, 37–39).

[Esp+12] Colin A. Espie, Simon D. Kyle, Chris Williams, et al. “A Randomized, Placebo-Controlled Trial of Online Cognitive Behavioral Therapy for Chronic Insomnia Disorder Delivered via an Automated Media-Rich Web Application”. In: *Sleep* 35.6 (2012), pp. 769–781. DOI: 10.5665/sleep.1872 (cit. on p. 11).

[Fog02] B. J. Fogg. “Persuasive technology”. In: *Ubiquity* 2002.December (Dec. 2002), p. 2. DOI: 10.1145/764008.763957 (cit. on p. 12).

[Fog03] B. J. Fogg. *Persuasive technology : using computers to change what we think and do*. Morgan Kaufmann Publishers, 2003, p. 283 (cit. on pp. 12, 13, 19, 20, 24, 66, 67).

[Fog09a] B. J. Fogg. “A behavior model for persuasive design”. In: *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive ’09*. 2009, p. 1. DOI: 10.1145/1541948.1541999 (cit. on pp. 12, 14).

[Fog09b] B. J. Fogg. “The Behavior Grid”. In: *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive ’09*. New York, New York, USA: ACM Press, 2009, p. 1. DOI: 10.1145/1541948.1542001 (cit. on pp. 12, 14, 15).

[Fog98] B. J. Fogg. “Persuasive computers”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’98*. New York, New York, USA: ACM Press, 1998, pp. 225–232. DOI: 10.1145/274644.274677 (cit. on p. 12).

[FT99] B. J. Fogg and Hsiang Tseng. “The elements of computer credibility”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI ’99*. 1999, pp. 80–87. DOI: 10.1145/302979.303001 (cit. on pp. 12, 20, 22).

[Hea+05] National Institutes of Health et al. “NIH state-of-the-science conference statement on manifestations and management of chronic insomnia in adults”. In: *NIH Consens Sci Statements* 22 (2005), pp. 1–30 (cit. on p. 2).

[Hea+06] National Heart, Blood Institute, National Institutes of Health, Us Department of Health, and Human Services. “Your Guide to Healthy Sleep”. In: (2006) (cit. on pp. 2, 3, 5).

[Hor+12] Corine Horsch, Willem-Paul Brinkman, Rogier Van Eijk, and Mark Neerincx. “Towards the usage of persuasive strategies in a virtual sleep coach”. In: *People, Computers & Psychotherapy* (2012) (cit. on p. 2).

[Jac+04] Gregg D. Jacobs, Edward F. Pace-Schott, Robert Stickgold, and Michael W. Otto. “Cognitive Behavior Therapy and Pharmacotherapy for Insomnia”. In: *Archives of Internal Medicine* 164.17 (2004), p. 1888. DOI: 10.1001/archinte.164.17.1888 (cit. on pp. 1, 2).

[JRB06] Eric O. Johnson, Thomas Roth, and Naomi Breslau. “The association of insomnia with anxiety disorders and depression: Exploration of the direction of risk”. In: *Journal of Psychiatric Research* 40.8 (Dec. 2006), pp. 700–708. DOI: 10.1016/J.JPSYCHIRES.2006.07.008 (cit. on p. 1).

[KD06] Spiro Kiousis and Daniela V Dimitrova. “Differential impact of Web site content: Exploring the influence of source (public relations versus news), modality, and participation on college students’ perceptions”. In: *Public Relations Review* 32 (2006), pp. 177–179. DOI: 10.1016/j.pubrev.2006.02.013 (cit. on p. 23).

[Kel+16] Saskia M. Kelders, Harri Oinas-Kukkonen, Anssi Oörni, and Julia E.W.C. van Gemert-Pijnen. “Health Behavior Change Support Systems as a research discipline; A viewpoint”. In: *International Journal of Medical Informatics* 96 (2016), pp. 3–10. DOI: 10.1016/j.ijmedinf.2016.06.022 (cit. on pp. 13, 16, 31).

[Kla15] Randy Klaassen. “HCI Perspectives on Behavior Change Support Systems”. SIKS dissertation series no. 2015-09. PhD thesis. University of Twente, Feb. 2015. DOI: 10.3990/1.9789036538398 (cit. on p. 13).

[Kuh+16] Eric Kuhn, Brandon J Weiss, Katherine L Taylor, et al. “CBT-I Coach: A Description and Clinician Perceptions of a Mobile App for Cognitive Behavioral Therapy for Insomnia.” In: *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 12.4 (Apr. 2016), pp. 597–606. DOI: 10.5664/jcsm.5700 (cit. on pp. 9, 10).

[Lee05] Teofilo Lee-Chiong. *Sleep: a comprehensive handbook*. John Wiley & Sons, 2005, p. 109. DOI: 10.1002/0471751723. arXiv: arXiv:1011.1669v3 (cit. on pp. 1, 3).

[LFE18] Annemarie I. Luik, Pedro Farias Machado, and Colin A. Espie. “Delivering digital cognitive behavioral therapy for insomnia at scale: does using a wearable device to estimate sleep influence therapy?” In: *npj Digital Medicine* 1.1 (Dec. 2018), p. 3. DOI: 10.1038/s41746-017-0010-4 (cit. on p. 3).

[LHS08] Bettina Laugwitz, Theo Held, and Martin Schrepp. “Construction and Evaluation of a User Experience Questionnaire”. In: *HCI and Usability for Education and Work* 5298 (2008), pp. 63–76. DOI: 10.1007/978-3-540-89350-9_6 (cit. on pp. 48, 86).

[LKC11] Chunsik Lee, Junga Kim, and Sylvia M Chan-Olmsted. “Branded product information search on the Web: The role of brand trust and credibility of online information sources”. In: *Journal of Marketing Communications* 17.5 (2011), pp. 355–374. DOI: 10.1080/13527266.2010.484128 (cit. on p. 23).

[LKE17] Annemarie I. Luik, Simon D. Kyle, and Colin A. Espie. “Digital Cognitive Behavioral Therapy (dCBT) for Insomnia: a State-of-the-Science Review”. In: *Current Sleep Medicine Reports* 3.2 (June 2017), pp. 48–56. DOI: 10.1007/s40675-017-0065-4 (cit. on pp. 5, 9).

[LM16] Yeonkyu Lee and Mikyung Moon. “Utilization and content evaluation of mobile applications for pregnancy, birth, and child care”. In: *Healthcare Informatics Research* 22.2 (Apr. 2016), pp. 73–80. DOI: 10.4258/hir.2016.22.2.73 (cit. on p. 23).

[LO11] Tuomas Lehto and Harri Oinas-Kukkonen. “Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature.” In: *Journal of medical Internet research* 13.3 (July 2011), e46. DOI: 10.2196/jmir.1559 (cit. on pp. 6, 17, 25, 31).

[Ltd18] QSR International Pty Ltd. *NVivo qualitative data analysis Software*. Version 12. 2018 (cit. on pp. 29, 54).

[MF08] Miriam J Metzger and Andrew J Flanagin. *Digital media, youth, and credibility*. Vol. 114. 2008, p. 202. DOI: 10.1162/dmal.9780262562324.vii (cit. on pp. 20, 21).

[Min+14] Yul Ha Min, Jong Won Lee, Yong Wook Shin, et al. “Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: A feasibility study”. In: *Journal of Medical Internet Research* 16.5 (May 2014), e135. DOI: 10.2196/jmir.3421 (cit. on p. 34).

[MY81] James C. McCroskey and Thomas J. Young. “Ethos and credibility: The construct and its measurement after three decades”. In: *Central States Speech Journal* 32.1 (1981), pp. 24–34. DOI: 10.1080/10510978109368075 (cit. on p. 23).

[OH09] Harri Oinas-Kukkonen and Marja Harjumaa. “Persuasive systems design: Key issues, process model, and system features”. In: *Communications of the Association for Information Systems* 24.1 (2009), pp. 485–500 (cit. on pp. 13, 14, 16–19, 21, 29, 64).

[Oha90] Roobina Ohanian. “Construction and validation of a scale to measure celebrity”. In: *Journal of Advertising* 19.3 (1990), pp. 39–52. DOI: 10.2307/4188769 (cit. on p. 23).

[Oin10] Harri Oinas-Kukkonen. “Behavior Change Support Systems: A Research Model and Agenda”. In: *Persuasive Technology*. Ed. by Thomas Ploug, Per Hasle, and Harri Oinas-Kukkonen. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 4–14 (cit. on p. 13).

[Pen+16] Wei Peng, Shaheen Kanthawala, Shupei Yuan, and Syed Ali Hussain. “A qualitative study of user perceptions of mobile health apps”. In: *BMC Public Health* 16.1 (Dec. 2016), p. 1158. DOI: 10.1186/s12889-016-3808-0 (cit. on p. 30).

[PGG92] S Mark Pancer, Margo George, and Robert J Gebotys. “Understanding and predicting attitudes towards computers”. In: *Computers in Human Behavior* 8.2-3 (1992), pp. 211–222 (cit. on p. 20).

[Rap10] Christof Rapp. “Aristotle’s Rhetoric”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2010. Metaphysics Research Lab, Stanford University, 2010 (cit. on p. 21).

[Rit+09] Lee M. Ritterband, Frances P. Thorndike, Linda A. Gonder-Frederick, et al. “Efficacy of an internet-based behavioral intervention for adults with insomnia”. In: *Archives of General Psychiatry* 66.7 (2009), pp. 692–698. DOI: 10.1001/archgenpsychiatry.2009.66 (cit. on p. 11).

[Rot07] Thomas Roth. “Insomnia: definition, prevalence, etiology, and consequences.” In: *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 3.5 Suppl (Aug. 2007), S7–10 (cit. on p. 1).

[SA12] Mahendra P Sharma and Chittaranjan Andrade. “Behavioral interventions for insomnia: Theory and practice”. In: *Indian Journal of Psychiatry* 54.4 (2012), p. 359. DOI: 10.4103/0019-5545.104825 (cit. on p. 2).

[Sad10] Sahoo Saddichha. “Diagnosis and treatment of chronic insomnia.” In: *Annals of Indian Academy of Neurology* 13.2 (Apr. 2010), pp. 94–102. DOI: 10.4103/0972-2327.64628 (cit. on p. 1).

[SHT14] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. “Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios”. In: *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*. Ed. by Aaron Marcus. Cham: Springer International Publishing, 2014, pp. 383–392. DOI: 10.1007/978-3-319-07668-3_37 (cit. on p. 48).

[SM03] James Brian Stiff and Paul a Mongeau. *Persuasive communication*. 2nd ed. Guilford Press, 2003, p. 351 (cit. on pp. 20, 23).

[SSR07] David Stewart, Prem Shamdasani, and Dennis Rook. "Analyzing Focus Group Data". In: *Focus Groups*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Ltd., 2007, pp. 109–133. DOI: 10.4135/9781412991841.d51 (cit. on p. 29).

[SST87] Arthur J Spielman, Paul Saskin, and Michael J Thorpy. "Treatment of Chronic Insomnia by Restriction of Time in Bed". In: *Sleep* 10.1 (1987), pp. 45–56 (cit. on p. 2).

[TF99] Shawn Tseng and B. J. Fogg. "Credibility and computing technology". In: *Communications of the ACM* 42.5 (May 1999), pp. 39–44. DOI: 10.1145/301353.301402 (cit. on pp. 12, 20).

[TO09] Kristian Torning and Harri Oinas-Kukkonen. "Persuasive system design : State of the art and future directions". In: *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*. New York, New York, USA: ACM Press, 2009, p. 1. DOI: 10.1145/1541948.1541989 (cit. on p. 26).

[Uyu+17] Begum Erten Uyumaz, Rosa Hendrikx, Laury Quaedackers, et al. "An interactive thought visualization tool for insomnia treatment". In: *Procedia Computer Science* 121 (Jan. 2017), pp. 314–321. DOI: 10.1016/J.PROCS.2017.11.043 (cit. on p. 32).

[VDM11] Laurian C Vega, Tom DeHart, and Enid Montague. *Trust between patients and health websites: A review of the literature and derived outcomes from empirical studies*. Nov. 2011. DOI: 10.1007/s12553-011-0010-3 (cit. on p. 23).

[Yu+18] Jessica S Yu, Eric Kuhn, Katherine E Miller, and Katherine Taylor. "Smartphone apps for insomnia: examining existing apps' usability and adherence to evidence-based principles for insomnia management". In: *Translational Behavioral Medicine* (Mar. 2018). DOI: 10.1093/tbm/iby014 (cit. on p. 6).

Webpages

[Ame17] American Academy of Sleep Medicine. *Two week sleep diary*. 2017. URL: http://sleepeducation.org/docs/default-document-library/sleep-diary.pdf (visited on Apr. 13, 2018) (cit. on p. 4).

[Goo18] Google Material Design. *Material Design Guidelines*. 2018. URL: https://material.io (visited on July 13, 2018) (cit. on pp. 36, 65).

[Kem] Kemepenhaeghe. *slaapgeneeskunde | Diagnostiek slaapstoornissen (EEG-/)slaapregistratie*. URL: https://www.kempenhaeghe.nl/download/13/13/161/0/0/Pati%7B%5C%22%7Be%7D%7Dnten-slaapgeneeskunde/ (visited on Apr. 13, 2018) (cit. on pp. 4, 77).

[Mar18] Christian Panadero Martinez. *Android Wizard*. Version 1.2. 2018. URL: https://github.com/PaNaVTEC/Wizard (visited on Aug. 20, 2018) (cit. on p. 44).

[Mic18] Microsoft. *Get Started with Android - Visual Studio App Center | Microsoft Docs*. 2018. URL: https://docs.microsoft.com/en-us/appcenter/sdk/getting-started/android (visited on Aug. 20, 2018) (cit. on p. 44).

[Pat] Pathosethoslogos.com. *Home - Ethos, Pathos, and Logos, the Modes of Persuasion Explanation and Examples*. URL: http://pathosethoslogos.com/ (visited on Mar. 5, 2018) (cit. on p. 21).

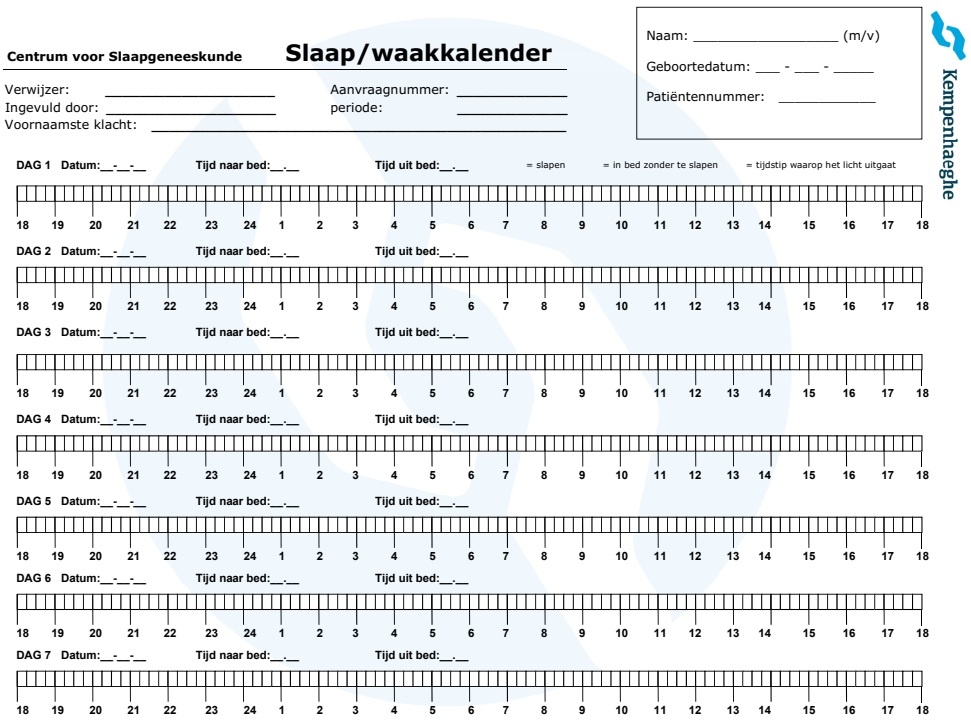
[Set11] Ramit Sethi. *The master of persuasion: Interview with BJ Fogg - Business Insider*. 2011. URL: https://www.businessinsider.com/the-master-of-persuasion-interview-with-bj-fogg-2011-2 (visited on Aug. 5, 2018) (cit. on p. 12).

[Sha] Dharmesh Shah. *The Importance of Google PageRank: A Guide For Small Business Executives*. URL: https://blog.hubspot.com/blog/tabid/6307/bid/45/the-importance-of-google-pagerank-a-guide-for-small-business-executives.aspx (visited on Mar. 5, 2018) (cit. on p. 19).

[Tuc] Tuck.com. *Insomnia - Definition, symptoms, treatments and more | Tuck Sleep*. URL: https://www.tuck.com/insomnia/ (visited on Mar. 5, 2018) (cit. on p. 1).

[Yar18] Yaroslav. *DiscreteScrollView*. 2018. URL: https://github.com/yarolegovich/DiscreteScrollView (visited on Aug. 20, 2018) (cit. on p. 44).

Sleep Diary - Kampenhaeghe



Centrum voor Slaapgeneeskunde **Slaap/waakkalender**

Verwijzer: _____ Aanvraagnummer: _____
Ingevuld door: _____ periode: _____
Voornaamste klacht: _____

Naam: _____ (m/v)
Geboortedatum: ___ - ___ - _____
Patiëntennummer: _____

Kempenhaeghe

DAG 1 **Datum:**__-__-__ **Tijd naar bed:**__.__ **Tijd uit bed:**__.__ = slapen = in bed zonder te slapen = tijdstip waarop het licht uitgaat

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

DAG 2 **Datum:**__-__-__ **Tijd naar bed:**__.__ **Tijd uit bed:**__.__

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

DAG 3 **Datum:**__-__-__ **Tijd naar bed:**__.__ **Tijd uit bed:**__.__

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

DAG 4 **Datum:**__-__-__ **Tijd naar bed:**__.__ **Tijd uit bed:**__.__

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

DAG 5 **Datum:**__-__-__ **Tijd naar bed:**__.__ **Tijd uit bed:**__.__

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

DAG 6 **Datum:**__-__-__ **Tijd naar bed:**__.__ **Tijd uit bed:**__.__

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

DAG 7 **Datum:**__-__-__ **Tijd naar bed:**__.__ **Tijd uit bed:**__.__

18 19 20 21 22 23 24 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Kempenhaeghe

OPMERKINGEN

DAG 1: _____

DAG 2: _____

DAG 3: _____

DAG 4: _____

DAG 5: _____

DAG 6: _____

DAG 7: _____

Invullen van de slaap/waakkalender

Instructie
Met deze kalender brengt u uw huidige slaappatroon in kaart. Dit is belangrijk om inzicht te krijgen in uw actuele slaapprobleem. Ook als u deze kalender al eens eerder hebt ingevuld, is het belangrijk om deze nu opnieuw in te vullen.

Het weekschema op de kalender loopt van dag 1 tot en met dag 7. De hele uren zijn aangegeven met een verticale streep en onderverdeeld in vier hokjes van elk een kwartier. De balk loopt van 's avonds 18.00 tot de volgende dag 18.00 uur.

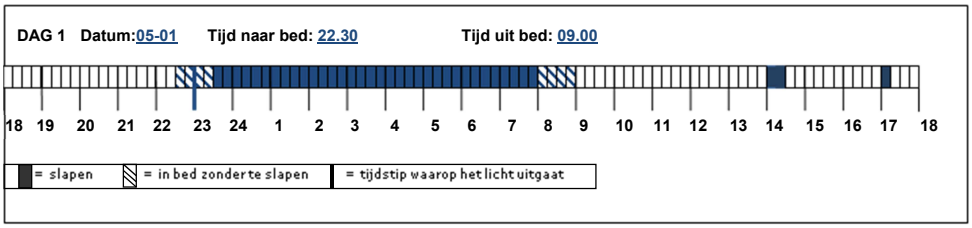
Belangrijk is dat u de kalender pas 's ochtends invult. Kijk 's nachts niet op de wekker; het gaat om een schatting.

De kalender vult u als volgt in:

- De tijd dat u slapend in bed doorbrengt, kleurt u geheel in
- De tijd die u in bed doorbracht zònder te slapen, geeft u aan door die vakjes te arceren (schuine streepjes).
- Het tijdstip waarop u het licht uitdoet, geeft u aan met een verticale streep.
- De tijd waarop u naar bed gaat, in slaap valt, wakker wordt en in bed ligt, of blijft liggen, rondt u af op kwartieren.
- Elk hokje staat voor 1 kwartier
- **Vul het schema in vanaf het moment dat u het vragenlijstenpakket binnen heeft en stuur de kalender samen met de vragenlijsten pas op als u het schema voor een volle week ingevuld heeft. Sla alstublieft geen dagen over.**

Ter verduidelijking een voorbeeld:

Op dag 1, 5 januari, bent u om 22.30 uur naar bed gegaan. Om 23.00 uur hebt u het licht uitgedaan, waarna u vervolgens om 23.30 in slaap bent gevallen. Op 6 januari werd u om 8.00 uur wakker, maar u bent nog een uurtje blijven liggen en om 9.00 uur opgestaan. Van 14.00 tot 14.30 uur bent u in slaap gevallen. Hetzelfde gebeurde om 17.00 uur. Deze laatste keer duurde de slaap echter naar tien minuten. Als u deze slaapperiodes in de kalender invult, ziet dat er als volgt uit:



Vragen?
Hebt u vragen over deze kalender? Neem dan gerust contact op met de slaapverpleegkundige in het Centrum voor slaapgeneeskunde, Locatie Heeze
T (040) 227 92 37 of locatie Oosterhout T (0162) 48 18 46, bgg (0162) 48 18 02
Versie juni 2016

Kempenhaeghe • Centrum voor Slaapgeneeskunde • Postbus 61 • 5590 AB Heeze • T (040) 227 90 22 / Postbus 250 • 4900 AG Oosterhout • T (0162) 481 800
W www.kempenhaeghe.nl • E info@kempenhaeghe.nl

Fig. A.1.: Sleep Diary developed by Kampenhaeghe
Source: Kampenhaeghe Patiënten-slaapgeneeskunde [Kem]

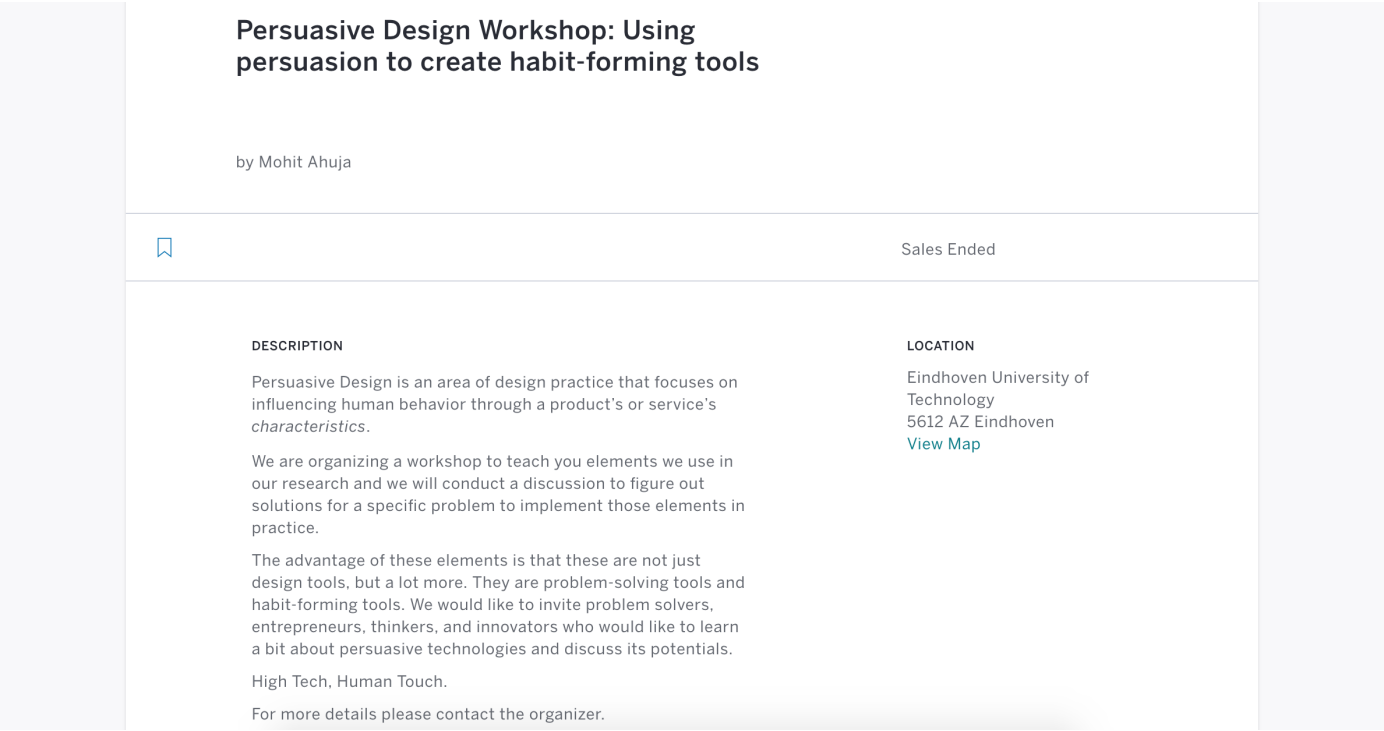


Fig. B.1.: Focus group event snapshot

Source: Event invitation can be found at https://habitformers.eventbrite.com

Focus Group Poster

Fig. C.1.: Focus Group Poster

Web App Snapshot

D

Fig. D.1.: Web app created for the workshop, a touch on a the design feature card flips it to show the definition of the design feature

Source: The tool can be found at https://bit.ly/psd-features

Negotiation Process

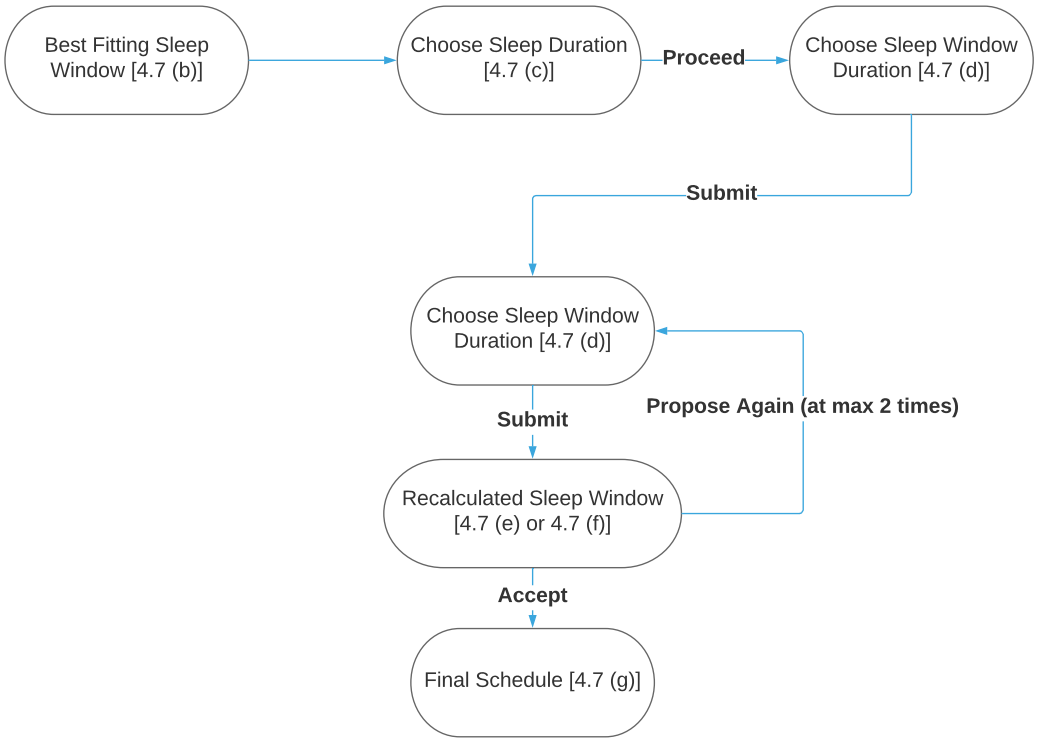


Fig. E.1.: Negotiation process for Sleep Scheduler, the corresponding screens can be found in Figure 4.7

Insomnia Severity Index

F

Part ID:_____

Sleep Behavior

Please rate the current severity of your sleep behavior.

Difficulty falling asleep.

None	Mild	Moderate	Severe	Very Severe
○	○	○	○	○

Difficulty staying asleep.

None	Mild	Moderate	Severe	Very Severe
○	○	○	○	○

Problem of waking up too early.

None	Mild	Moderate	Severe	Very Severe
○	○	○	○	○

How SATISFIED/DISSATISFIED are you with your CURRENT sleep pattern?

Very Satisfied	Satisfied	Moderately Satisfied	Dissatisfied	Very Dissatisfied
○	○	○	○	○

How NOTICEABLE to others do you think your sleep problem is in terms of impairing the quality of your life?

Not at all Noticeable	A Little	Somewhat	Much	Very Much Noticeable
○	○	○	○	○

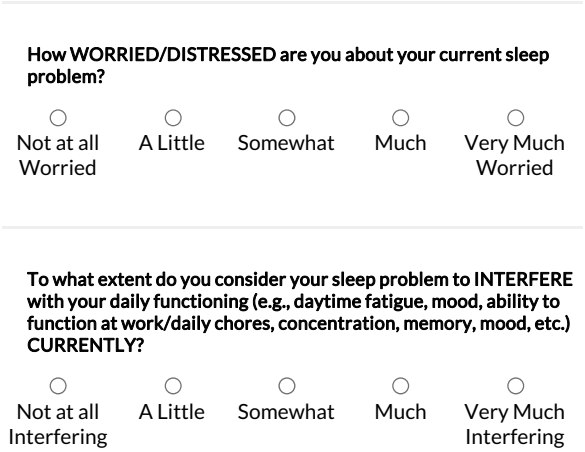


Fig. F.1.: Insomnia Severity Index developed by Bastien et al.
Source: [BVM01]

Credibility Expectancy Questionnaire

G

Therapy Evaluation Form

We would like you to indicate below how much you believe, right now, that the therapy you are receiving will help to improve your lifestyle / functioning. Belief usually has two aspects to it: (1) what one thinks will happen and (2) what one feels will happen. Sometimes these are similar; sometimes they are different. Please answer the questions below. In the first set, answer in terms of what you think. In the second set answer in terms of what you really and truly feel. We do not want your course convenors to ever see these ratings, so please keep the sheet covered when you are done.

Set I

1. At this point, how logical does the course offered to you seem?

1	2	3	4	5	6	7	8	9
not at all logical				somewhat logical				very logical

2. At this point, how successfully do you think this course will be in raising the quality of your functioning?

1	2	3	4	5	6	7	8	9
not at all useful				somewhat useful				very useful

3. How confident would you be in recommending this course to a friend who experiences similar problems?

1	2	3	4	5	6	7	8	9
none at all confident				Somewhat confident				very confident

4. By the end of the course, how much improvement in your functioning do you think will occur?

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Set II

For this set, close your eyes for a few moments, and try to identify what you really feel about the course and its likely success. Then answer the following questions.

1. At this point, how much do you really feel that the course will help you to improve your functioning?

1	2	3	4	5	6	7	8	9
not at all				somewhat				very much

2. By the end of the course, how much improvement in your functioning do you really feel will occur?

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Fig. G.1.: Credibility Expectancy Questionnaire developed by Devilly et al.
Source: [DB00]

User Experience Questionnaire

H

Please make your evaluation now.

For the assessment of the product, please fill out the following questionnaire. The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression.

Example:

attractive	○	⊗	○	○	○	○	○	unattractive

This response would mean that you rate the application as more attractive than unattractive.

Please decide spontaneously. Don't think too long about your decision to make sure that you convey your original impression.

Sometimes you may not be completely sure about your agreement with a particular attribute or you may find that the attribute does not apply completely to the particular product. Nevertheless, please tick a circle in every line.

It is your personal opinion that counts. Please remember: there is no wrong or right answer!

Please assess the product now by ticking one circle per line.

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Fig. H.1.: User Experience Questionnaire developed by Laugwitz et al.
Source: [LHS08]

Complete Thematic Analysis for focus group

Tab. I.1.: Thematic Analysis of Focus Group Discussions complete

Theme	SubTheme (Coverage)	Description	Example Quotes
Primary Task	Personalization (13.89%)	Customise the content or services to a user's level	*You can receive a remark do you want to do it this way, or you can do it in a better way.* *If you can personalise the app. Then everybody can choose what puts them to sleep. You can't have a standard way like music put everybody to sleep. Maybe it puts me to sleep doesn't put them to sleep. It might wake somebody else up.* *I would prefer to have personalised advice.*
	Reduction (6.15%)	Simplify the task the users are trying to do.	*The way I would think about it is I'm trying to get somebody to start the habit, so, I will try to find the smallest thing that he can do, that it's easy for him to do.* *So, make something that is easy and makes it easier for them to keep repeating.*
	Simulation (5.14%)	Systems should provide simulations so they can immediately observe the link between cause and effect.	*For example, I'm using this calorie thingie app, like where you count galleries every day. You know, I am super motivated to maintain the calories. So if I go overboard then I know I can't do this. So you can create an app that which actually monitors your sleep, or you can feed in for how many hours you asleep and then if you like to build in some features where you tell this person that if you sleep so much, you have so many benefits.*

Continued on next page

Tab.I.1 – continued from previous page

Theme	SubTheme (Coverage)	Description	Example Quotes
	Tailoring (4.05%)	Tailor the system for the users as part of a group suiting to their potential needs, interests, personality, usage context, or other factors relevant to them.	*It should be adjusted to both; sometimes you are active or lazy. And also customise should be more important because people have different reasons for not sleeping on time.* *Very intensive coaching or very light coaching? If you choose the light way, you can build so much stimulus so that in a month you are involved more, and the stimulus takes from low interaction to intensive interaction in a month time. Because they get part by part more interested. Keep them ingeniously learning about themselves.*
	Tunneling (1.94%)	Guide the user step-by-step through a (new) process.	*so it's just like the show that you are overboard or a bit negative from your target calories so that you can exercise more and burn something. Something like that, not very directly but somewhere hinting about the further process.*
	Self-monitoring (1.34%)	Enable them to track their progress to alter behaviour to achieve a pre-meditated outcome	*maybe create an app where you can see as an energy meter for your body or something so mainly displays how it understands you better. Shows your energy levels.*
Dialogue	Liking (6.48%)	The aesthetic attractiveness of a user to its audience makes it be perceived as persuasive	*They shouldn't find it tedious to not even use it at all.* *l kids around two to eight years, somewhere like that and for me probably because I'm working with robots you know what I would want to do it maybe I can develop a sleeping companion for them a little sleeping part that they can cuddle when they are sleeping.*
	Similarity (4.93%)	People look at the system, and the system should imitate them in certain ways.	*You can couple it with the lighting in your house. So that at eleven o clock light will go off.* *The system should be able to identify if I am having more coffee in the morning that I wasn't able to sleep well last night, or if I have a smiley face then it is more likely I had good sleep.*

Continued on next page

Tab.I.1 – continued from previous page

Theme	SubTheme (Coverage)	Description	Example Quotes
	Rewards (4.84%)	A system giving its user virtual (or real credits), the users will more likely achieve their goals	*Specially in the beginning, the progress will be really slow. In the beginning, you need an external reward that's not just improved behaviour. So, in the beginning, you need something extra, some other rewards outside cannot be just sleeping better.* *you already identified that you're personal problem is that you are behind the computer or on your phone too long before you go to bed, and you want to do something about it, you put the phone/screen away then you get rewards.* *yes so they could say, yeah tonight you didn't check your phone three hours before you went to sleep, great, you get three points.*
	Suggestion (3.17%)	Systems offering fitting suggestions based on ongoing use have greater persuasion powers	*If this system said okay, I know that you drink based on your skin feedback that you drink four coffee between four pm to eight pm parade so no wonder you have too much caffeine that you don't fall into a deep sleep.* *customised suggestion of monitoring or coaching of their behaviour.* *use the data differently to give different suggestions.*
	Praise (2.31%)	Give (positive) feedback to the user using images/ symbols/ sounds/ words	*Engage people to maybe insist that habit. The idea in itself is that the occasion can be made more interesting.*
	Social Roles (1.59%)	When a system takes a social role when giving feedback, the system increases its persuasive capabilities	*If there is something wrong that is an unfortunate result of my sleep cycle then the application can ask what happened yesterday, or were you nervous? like that it can get the information from the user in some easier way?*
	Reminders (1.45%)	A system reminding its users of target behaviour in an ideal amount has the propensity to help its user achieve their goals	*So, make something that is easy and keeps telling them at a specific fixed time and making it easier for them to keep repeating.*

Continued on next page

Tab.I.1 – continued from previous page

Theme	SubTheme (Coverage)	Description	Example Quotes
Credibility	Trustworthiness (2.96%)	A trustworthy looking system has increased capability of persuasion as well	*But I think it's also good to know some general things.* *There are hospitals which work on sleep disorders and all. So I can ask that person to go to some trustful source (using the feature trustworthiness), and that maybe imbibes confidence in them that this is a serious issue because me telling them that won't have so much of the effect to it as a doctor specialised in sleep.*
	Authority (2.18%)	A system that leverages roles of authority will have enhanced powers of persuasion	*They go to a doctor who sends them to a sleep centre where you are excluded from the external influence which disrupts your sleep.* *Probably, you can look at scenarios or things which have a negative impact on the sleep. So you can think of creating virtual space.* *i would tell them to go to a doctor.* *I would motivate them to go to an expert.*
	Surface Credibility (2.12%)	People make an initial assessment of the system is credible or not based on a first-hand inspection., so a system with a competent look and feel, tend to be more persuasive	*for many years I could not do any practice or any movement, but later on, I realised it is getting worse and when I started to work.* *Then I joined a sports club, and I started doing yoga courses. Yoga now stretches your body, and it has the same movements that the doctor gave me. I began to enjoy, also being in that class. I also started to gain some results*
	Expertise (1.76%)	A system perceived as incorporating expertise, knowledge, and competence will have increased powers of persuasion. e.g. help features on most of the modern apps	*If your app can sometimes tell that if you don't sleep 30 hours in a week, then something terrible will happen to you.* *may be given them a personal life example that this has changed.*
	Verifiability (1.03%)	The system should provide verifiable information which they claim by giving a reference where they can be cross-checked	*tell them in some funnier exaggerated way.* *Give some medical evidence.*
	Third Party Endorsement (0.55%)	Third-party endorsements and recommendations from other trusted entities counts	*a partnership with a company which has an app where people are already recording what food they eat.*

Continued on next page

Tab.I.1 – continued from previous page

Theme	SubTheme (Coverage)	Description	Example Quotes
Social	Cooperation (9.23%)	A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to co-operate; thus it should provide means for co-operation with others	*If you make a plan for today for sleep you could send your sleep plan or the time you go to bed to your friends and to make your friend help you to land or get to sleep.* *Maybe you can use as an app coupling thing. This is based on the idea that I care more for the people I love than I do myself. So, if you couple this app to someone I love or like or respect I will probably love seeing their data than my data so giving someone in authority to go through my sleep data may help.* *i can use social cooperation. If I know someone is not doing well. I can say I can help you do well. Motivate each other. You wake up, and you wake me up.*
	Normative Influence (3.75%)	A system can leverage normative influence (peer pressure), thus provide means for them to see the norm, it influences their decision to increase the likelihood to adopt a target behaviour	*You still may need to find a social network to support this. If that's a three or four of you, drink too much alcohol in the evening, or you are too much awake. Then you are in a social context to influence each other.* *tell them something socially undesirable event will take place if they don't follow through.*
	Social Facilitation (2.76%)	Users are more likely to perform target behaviour if they can find via the system that others are performing the same behaviour along with them. Thus, a system should provide the means of finding others performing the same target behaviour	*If it is not your problem, or your partner, your child, or your mother or neighbour. If they have a good social network, then you might be able to have a different role in this.* *It is the social interaction which helps.*
	Social Learning (1.68%)	People look at each other while performing behaviours, so allow them to look at their social network who are trying to perform the same target behaviour to motivate them	*you can create an app where people share what get them to sleep and share it what can get them sleep and share their ideas of getting sleep. Sharing this will help others try this instead of the medicine.*
	Competition (0.67%)	A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to compete	*Sleep competition.* *You can compete with yourselves as well, I slept better than my average today, that's amazing.*

Continued on next page

Tab.I.1 – continued from previous page

Theme	SubTheme (Coverage)	Description	Example Quotes
	Social Comparison (0.37%)	System should provide users with means to compare performance with others, that will have greater motivation for them to perform the target behaviour	*You can compare yourself with how you stand with your friends*

Declaration

I hereby declare that I have completed this work solely and only with the help of references as mentioned. The work is carried out in Philips, Eindhoven and TU/e, Eindhoven.

Enschede, the Netherlands, September 24, 2018

Mohit Ahuja