# Understanding User-perception of Sleep to Inform Sensing and Provide Actionable Feedback

Ruth Ravichandran

A dissertationsubmitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Shwetak Patel, Chair

Julie A. Kientz

Payman Arabshahi

Program Authorized to Offer Degree:
Electrical and Computer Engineering

University of Washington

## Abstract

# Understanding User-perception of Sleep to Inform Sensing and Provide Actionable Feedback

Ruth Ravichandran

Chair of the Supervisory Committee:
Professor Shwetak Patel
Allen School of Computer Science & Engineering
Department of Electrical & Computer Engineering

It is becoming increasingly clear of the importance of sleep and its impact on our daily lives. Despite the pervasiveness of sleep issues, people struggle to assess and improve their sleep. Sleep is an unconscious, passive activity and accurate, manual, self-tracking of sleep is often unattainable. A majority of the general population remains unaware of long-term patterns in their sleep duration and consistency, unless the individual suffers from a chronic sleep related disorder that affects day-time functioning or causes other related health issues. The growing popularity of commercial sleep sensors for use at home shows that these technologies have the potential to provide long-term, low-cost, and accurate representations of people's daily sleep patterns in the comfort of their home environment. However, current sleep tracking devices are limited by hardware and algorithmic constraints. Unactionable sleep feedback, in the context of long term sleep tracking for individuals without sleep disorders, leads to mental models that are in tension with recommendations for good sleep health from sleep medicine professionals. Through a comprehensive survey of commercial sleep sensing devices, I identify gaps between user perception of sleep data and best practices in sleep health. I contribute to the design, implementation and evaluation of a wireless sleep sensing system that enables non-contact monitoring of sleep related physiological signals throughout the night. I also explore, through the design and implementation of a smart phone app, how a personalized sleep model can be used to provide meaningful and actionable feedback that can provide users with the intent to make changes to their behaviour to improve sleep health.

# TABLE OF CONTENTS

# LIST OF FIGURES

# GLOSSARY

ACTIGRAPHY: A non-invasive method of monitoring human rest/activity cycles.

DOPPLER: A specialized radar that uses the Doppler effect to produce velocity data about objects at a distance. It does this by bouncing a microwave signal off a desired target and analyzing how the object's motion has altered the frequency of the returned signal.

EMA: Ecological momentary assessment (EMA) involves repeated sampling of subjects' current behaviors and experiences in real time, in subjects' natural environments.

N-BACK: The n-back task is a continuous performance task that is commonly used as an assessment in cognitive neuroscience to measure a part of working memory and working memory capacity. The subject is presented with a sequence of stimuli, and the task consists of indicating when the current stimulus matches the one from n steps earlier in the sequence. The load factor n can be adjusted to make the task more or less difficult.

PSG: Polysomnography, a type of sleep study, is a multi-parametric test used in the study of sleep and as a diagnostic tool in sleep medicine.

PVT: The psychomotor vigilance task (PVT) is a sustained-attention, reaction-timed task that measures the speed with which subjects respond to a visual stimulus. Research indicates increased sleep debt or sleep deficit correlates with deteriorated alertness, slower problem-solving, declined psycho-motor skills, and increased rate of false responding.

RADAR: Radio Detection and Ranging is a detection system that uses radio waves to determine the range, angle, or velocity of objects.

# ABBREVIATIONS

BR: Breathing Rate

EMA: Ecological Momentary Assessment

HR: Heart Rate

RADAR: Radio Detection and Ranging

CPM: Cycles Per Minute

PCA: Principal Component Analysis

PSG: Polysomnography

REM: Rapid Eye Movement

PVT: Psychomotor Vigilance Test

# ACKNOWLEDGMENTS

I thank the Lord Almighty for His continued grace. This research work would not have possible without the guidance and support of my advisor Shwetak Patel and co-advisor Julie Kientz. I am incredibly grateful and indebted to you for the opportunity to work in your lab alongside some of the best researchers in UbiComp and HCI. These past years that I got to work in the UbiComp lab were life-changing. You created an environment that fostered collaboration and gave us the freedom to explore different research avenues without restraint. Shwetak, thank you for supporting me at my weakest moments and for never running out of ideas and ways of making things better. I hope to one day see in myself what you saw in me. Julie, you have been an inspiration and a role model to me as a researcher and a mother. I would not have survived graduate school without your mentorship. I have always looked up to you for your extraordinary work ethic and seemingly endless capacity for kindness. Thank you for opening your arms to me when I needed comfort during some of the hardest moments in my personal life.

To my seniors and mentors from the UbiComp lab, Gabe, Sidhant, Mayank and Keyu, you set the bar high with your work and contributions to the UbiComp community. Gabe, thank you for giving me the first opportunity to work in the UbiComp lab on the static e-field sensing project. I got to see first hand how to you approach problems and come up with sound scientific explanations for seemingly bizarre and unknown behaviours. Sidhant, thank you being a friend and an ever present mentor during your time in the lab. Thank you for spending an entire summer helping me figure out GNU radios and USRPS. You presented me with the greatest opportunity of working on WiBreathe, and that set the trajectory for the rest of my Phd. For that I will always be grateful to you. Mayank, you made our lab such a fun and stress free place. You are a bottomless pit of enthusiasm. Thank you for holding my hand through my first publication. Keyu, you are the older brother I never had. Your wise words and insights at critical moments during graduate school sustained me and provided much needed encouragement and motivation. I hope to one day live up to your standards of perfectionism in work. You helped me get an amazing job offer and you probably changed the course of my life, for that I am indebted to you.

My lab mates at the UbiComp lab, I couldn't have asked for a better lab to work at. Just working alongside you has changed me for the better in so many ways. I have passively absorbed traits from each and everyone of you. Elliot, what would I have done without you! Thank for being such a amazing friend, a never ending source of knowledge and information for my incessant questions and for being a great partner in all the EM courses. Your love for all things signal processing was infectious and I may have caught some of it too. I cherish all the deep philosophical, moral and spiritual conversations we had over bread bowls at the HUB. Edward, I admire your tenacity and capacity to organize and execute multiple projects at once. Over conversations on high level research questions were always insightful. Alex, thank you for you the numerous times your chipped in to rescued me during deadlines and before presentations. I am forever grateful for all the feedback you gave on my research talks and the statistics help. Lilian, thank you for supporting me through a tough phase in my life. Our weekly workouts carried me through those difficult phases. Hanchuan, Mohit and Josh and Tien thank your the valuable insights your shared through seemingly random conversations and small talk. Eric, thank you to for allowing me to ride alongside you during the ideation phase of your projects. It was fun to see you tackle problems with such vigor and enthusiasm. I hope I can be as good at debugging as you are someday.

To all my wonderful collaborators, thank you for supporting me and working along side me in the projects that made this thesis possible. Rifat, thank you for extending me the opportunity to collaborate with you on DoppleSleep. It has given me incredible amounts of exposure in the research community. Laura, thank you for giving me a crash course in how to do qualitative research and working with me closely on Making Sense of Sleep Sensors. Your endurance and focus helped us get to the finish line.

I would also like to acknowledge all the wonderful professors in the EE department who shared their knowledge and guided me through scientific advancement. I specially want to thank Prof. Denise Wilson who first seeded in me the idea to pursue graduate school. Thank you for the for believing in me and giving me this life-changing opportunity.

Last but not the least, I would like to acknowledge the enduring support I received from my family members. Ebenezer, thank you for encouraging me to pursue my PhD and for always being ,my source of strength. I could depend on you at all times and I wouldn’t have accomplished this without you. Mom and dad thank you for all the sacrifices you made to give me a strong foundational early education. Thank you for instilling in me the confidence to be who I am.

# DEDICATION

to my dearest daughter, Heera

# Chapter 1

# SLEEP AS A MEASURE OF WELL-BEING

Good sleep is essential to good health. Adequate, restful sleep is as important to one's well-being as a healthy diet and regular physical activity. During sleep, the body and brain undergo necessary restorative activities [116], and inadequate sleep leads to reduced alertness and drowsiness [55]. In the United States, an estimated of 50 million people have poor sleep quality or have a sleep disorder such as insomnia, sleep apnea, and narcolepsy [34]. The average American worker reported 5.3 days of difficulty falling asleep, 6.6 days of trouble staying asleep, and 5.0 days of trouble waking up for work in the past month. Across the three types of poor sleep quality, work overload was positively associated with the frequency of poor sleep quality [73].

## *1.1 Motivation*

Despite the pervasiveness of sleep issues, people struggle to assess and improve their sleep. The reason is that sleep is an unconscious, passive activity and therefore, unlike diet and physical activity, which are difficult but possible to track manually [37], accurately self-tracking sleep manually is often unattainable. A majority of the general population remains unaware of their overall sleep quality or longer-term patterns in their sleep duration and consistency [119]. Unless the individual suffers from a chronic sleep related disorder that affects day-time functioning or causes other related health issues, they are unaware of their overall quality and it's effects.

When commercial fitness trackers were introduced around the year 2008, their primary goal was to enable people to keep track of the number of steps taken over the course of day which could then be used to get an estimate of the number of calories burned. The instant feedback that was made available to not only brought awareness but also enabled goal setting accompanied by behaviour change in users [71, 93]

In addition to step tracking, wearable devices further integrated several functionalities like heart rate monitoring, and sleep tracking functionality, since the sensors that tracked movement could also be used to track the lack of movement associated with sleep. About 172 million fitness tracking devices were shipped in the year 2018 showing their growing popularity. Despite the widespread adoption of these devices, their potential to facilitate meaningful behaviour has not been fully realized [89].

This has been reflected in the findings from several studies done by the HCI community exploring the different facets of user interaction with tracking devices. Some of the major issues that have been highlighted in these studies are: users' lack of trust in device data, broken mental model surrounding accuracy of the devices and hardware limitations [122]. Liang et. al, [88] have explored how the lack of accuracy in sleep tracking devices affects the users' perception of the feedback and thereby it's usefulness. Specifically, how the lack of definition of sleep metrics, limitations in underlying data collection and processing mechanisms, and lack of rigor in tracking approach causes errors in the feedback given to the users. While the findings from these studies shed light on some of the technological and algorithmic issues with sleep tracking devices, the question of how long term sleep tracking can facilitate improvement in sleep health is still unanswered. Long term sleep tracking for people with no clinically diagnosed sleep disorders, has some unique challenges.

Unlike diet and physical activity, for which users can set short term goals and use feedback from their health tracking devices to make behaviour changes, goal setting for improving sleep health may not always have immediately obvious measures and the effects of poor sleep quality drastically vary from person to person [26]. While the National Sleep Foundation has published population level recommendations for the duration of sleep for different age groups [60], studies have shown individuals' sleep varies by only as few as 23 minutes nightly over the course of the year [74]. Therefore sleep feedback focused on objective measures such as sleep duration and sleep efficiency alone may not provide users with the intention to make behaviour changes that promote good sleep health. The feedback from current commercial sleep sensing devices, intended for long-term sleep tracking, largely tend to mimic objective measures from clinical reports used for the diagnosis of sleep disorders carried out over over a short period of time. The overarching question still remains: how can sleep sensing devices used for long-term sleep tracking fully realize their potential to promote good sleep health? This leads to my thesis statement:

## *1.2 Thesis Statement*

Current sleep tracking devices are limited by hardware and algorithmic constraints. Unactionable sleep feedback, in the context of long term sleep tracking for individuals without sleep disorders, leads to mental models that are in tension with recommendations for good sleep health from sleep medicine professionals. The goal of this dissertation is to provide support for my thesis statement:

> Understanding the gaps between user perception of sleep data and expert opinions on good sleep practices can inform the design of novel sleep sensors and the development of personalized sleep models that can provide users with actionable feedback to improve their sleep health.

## *1.3 Research Questions and Approaches*

In my research, I have attempted to support my thesis statement by answering the following research questions.

- RQ1: How do users perceive the feedback they receive from current commercially available sleep sensing devices?

To understand user perception of sleep data I conducted a comprehensive survey of commercial sleep sensing devices through amazon reviews, android and apple app store reviews of popular sleep tracking devices. I supplemented this data with interviews and survey data from long term users of sleep tracking devices. In order to find the gaps between user perception of sleep data and best practices in sleep health, I conducted interviews with medical professionals specializing in sleep health.

The first finding from the study was that current sleep sensing devices act as facilitators to good sleep health by allowing users to monitor sleep in the comfort of their own homes, thereby bringing awareness to the users about the long term trends in sleep quality. However, current devices are limited by their battery life, invasiveness and range. Users find it uncomfortable to wear a device on their arm, or head while they sleep and sometimes forget to recharge the tracker they've been wearing all day.

The second finding from the study was that sleep related feedback provided by a majority of sleep sensing devices tend to focus on sleep scores and metrics that are derived from diagnostic measures used for clinical purposes. They do not take into account inter-individual trait variability in sleep variables and the effects of sleep. This detracts users from making behavior changes that may improve sleep health. The findings from the exploratory study raised two subsequent research questions.

- RQ2: How can we enable long-term sleep tracking in users' homes in an unobtrusive manner?

In order to address the above research question I have explored how sleep related physiological bio-markers can be tracked unobtrusively, at a distance, using RF signals. Radio frequency signals in the k-band range can be used to sense minute chest movements due to breathing and heart beats in addition to coarse body movements. Features extracted from these sleep related physiological bio-markers can be fed into a machine learning model to better classify sleep related metrics.

- RQ3: How can we provide actionable sleep feedback that can help users adopt healthy sleep habits?

Previous work has shown that providing timely, contextual and actionable feedback can help people make behavioral changes to improve their health. In the context of sleep tracking, this extends beyond tracking night time sleep to tracking the different variables that affect sleep as well as the effects of sleep on day time functioning. In order to facilitate this, I developed an smartphone app: SleepApp, that allows users to track their pre-bedtime routines as well as their cognitive function during the day in addition to their night time sleep. This data can then be used to develop a personalized sleep model that can provide users with actionable feedback to help improve their sleep health.

### *1.4 A Summary of Research Questions and Approaches*

I have summarized the research questions and approaches I used to answer them in Table 1.1.

Table 1.1: Studies conducted to answer the research questions

| **Research Questions** | **Studies Conducted** |
|---|---|
| How do users perceive the feedback they receive from current commercially available sleep sensing devices? | Making Sense of Sleep Sensors |
| How can we enable long-term sleep tracking in users' homes in an unobtrusive manner? | WiBreathe<br>DoppleSleep |
| How can we provide actionable sleep feedback that can help users adopt healthy sleep habits? | SleepApp |

## *1.5 Contributions*

In this dissertation, I present three types of contributions: artifact, empirical finding and methodology.

- In the first study I explore gaps in user perception of sleep data and provide guidelines for long-term everyday sleep tracking by adapting from best practices in sleep medicine. These guidelines include 1) focusing on actionable data rather than everyday scores 2) paying attention to long term trends in sleep quality and 3) accommodating for imprecise sleep metric estimations caused by hardware limitations in reported feedback.

- Based on the findings from the first study I contribute to the design, implementation and evaluation of a 2.4GHZ continuous wave wireless system that enables non-contact breathing rate estimation and the evaluation of a 5GHz continuous wave Doppler system that enables non-contact monitoring of sleep related physiological signals throughout the night. These systems overcome limitations in current commercial sleep sensing devices by allowing for longer sensing range, minimal intrusiveness and higher levels of accuracy.

- I also contribute to the design, implementation and evaluation of a smart phone app that facilitates tracking of variables and activities that affect sleep quality and by extension the effects on day-time cognitive function. A personalized model built from the data collected using the app over an extended period of time can be used to provide meaningful correlations between the variables that affect sleep and daytime cognitive function. Through findings from a follow-up survey I have shown how contextual and actionable feedback may give users the intention to make behaviour change.

### *1.6 Dissertation Overview*

The document is organized as follows:

- Chapter 2 provides the background for sleep health and sleep sensing. It covers the physiology of sleep,and the various metrics used to describe and quantify sleep quality and sleep health. It summarizes the different methods used for sleep tracking from the most accurate and highly invasive to the less accurate and invasive methods. The chapter also provides a background of the studies done the HCI community on long-term sleep tracking, user perception of sleep feedback and self-experimentation for sleep.

- Chapter 3 outlines a study that surveys how feedback provided by currently available commercial sleep sensing devices can sometimes act as a barrier to sleep health. This chapter highlights how users' perceptions of sleep feedback from these devices leads to broken mental model that are in tension with evidence-based methods for promoting good sleep health. This work was published at CHI'17.

- Chapter 4 presents WiBreathe, an initial exploration of continuous and non-contact physiological signal monitoring using continuous wave wireless signals in the ISM band. This work was published at IEEE PerCom 2015.

- Chapter 5 describes DoppleSleep, a system that examines the feasibility of non-contact sleep monitoring by tracking three significant sleep-related bio-markers: breathing rate, heart rate, and body motion using a K-band continuous wave Doppler RADAR wireless signals. This chapters outlines how a machine learning model based on the features extracted from the sleep related physiological bio-markers can be used to estimate sleep metrics with higher levels of accuracy compared to current devices while being minimally invasive to the users. This work was done in collaboration with Tauhidur Rahman et al from Cornell University and published at UbiComp'15.

- Chapter 6 provides guidelines for a smartphone based application that can help generate meaningful sleep related feedback to promote behaviour change. This chapters provides details of an approach that can be a promising, holistic measure of sleep health that is personally contextualized and can serve as a basis for meaningful self-experimentation. This chapter also contains a case-study of how this method can be used to understand sleep habits of the shift-worker population.

- Chapter 7 discusses the limitations of my present work and sets up opportunities for future work. I also discuss lessons learned during the course of my research.

- Chapter 8 summarizes and concludes this body of work.

Figure 1.1 shows the overview of my dissertation.

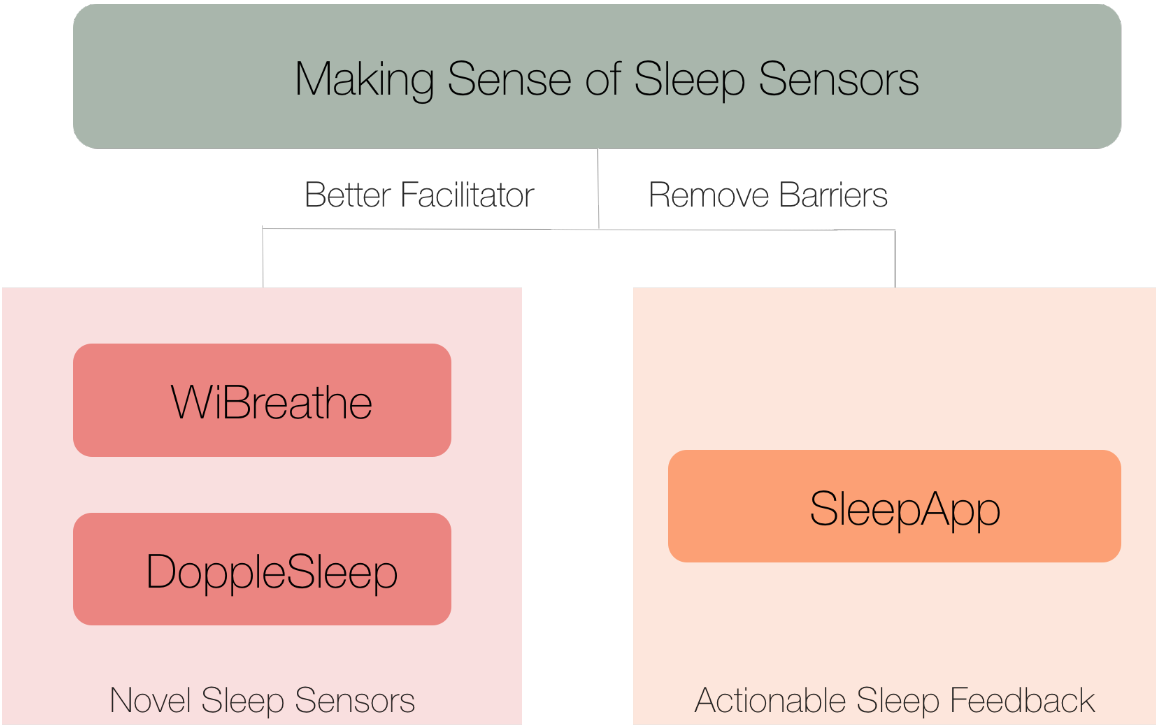


Figure 1.1: Dissertation scope. Study design of Making Sense of Sleep Sensors is described in Chapter 3. Design and evaluation of WiBreathe is described in Chapter 4. DoppleSleep design and implementation is described in Chapter 5 and SleepApp study design is described in Chapter 6.

## Chapter 2

## BACKGROUND AND RELATED WORK

The focus of this dissertation is the exploration of facets of sleep monitoring from a novel sensing perspective to providing feedback that promotes well-being. In this section, I have provided an overview of the physiology of sleep , the different ways in which sleep can be sensed, the metrics used to define sleep, sleep hygiene and other variables that affect asleep, novel computing based interventions for long term sleep tracking and the studies that look into supporting behaviour change through sleep self-experimentation. This large body of related work provides an understanding of how to develop sleep technology to support overall health and provide personalized recommendations based on both objective and subjective data. I have also situated my contribution in the literature on sleep tracking and self-experimentation and describe how my our work builds upon and extends the this body of work.

### *2.1 Physiology of Sleep*

Consuming almost a third of our daily lives, sleep is a significant marker of an individual's health and well-being. Getting ample, good-quality sleep facilitates higher levels of productivity, mental performance, and physical growth [63]. In order to understand sleep as a measurable entity and one that can be sensed or tracked, we need to understand the physiological changes that the body goes through during sleep. Understanding this can shed light on some of the metrics that will be used frequently in the paper and give better insight into the tensions that will be highlighted in the following sections. Typically, there are two main stages through which we cycle through when we sleep: Rapid Eye Movement (REM) sleep and three stages of Non Rapid Eye Movement (NREM) sleep labelled as N1, N2 and N3. as we transition form wakefulness to sleep and through the different stages, physiological signals such as the intensity of body motion, breathing rate and heart rate also vary corresponding to the various sleep stages [72].

**NREM** - 75% of our sleep is comprised of NREM sleep and is characterized by these changes in physiological signals:

**NREM1** - Also known as light sleep, is between being awake and falling asleep. This stage is characterized by cessation in muscle movement.

**NREM2** - It's the onset of sleep and breathing and heart rate becomes regular and body temperature drops. A deeper sleep where your body temperature cools a little and you become disengaged from your surroundings.

**NREM3** - Is the deepest and most restorative sleep. Blood pressure drops, breathing becomes slower and muscles are relaxed. It is very hard to wake up from deep sleep because this is when there is the lowest amount of activity in your body. It is also the part of sleep where your body rebuilds itself, restores energy and hormones are released.

**REM** - The remaining 25% of sleep falls into REM. It occurs about 90 minutes after falling asleep and recurs about every 90 minutes, getting longer later in the night. The body becomes immobile and relaxed. This stage characterized by rapid eye movement is also marked by rapid, irregular and shallow breathing. The heart rate and blood pressure increase.

## *2.2 Sensing Sleep*

During various stages of sleep, the body exhibits several bio-markers as described in the section above, therefore the onset of sleep and transitions between stages can be measured by tracking these sleep related bio-markers. Sensing sleep by measuring the physiological changes associated with sleep can help quantify the quality of sleep and diagnose disorders related to it.

### *2.2.1 Polysomnography*

Some people suspected of having sleep disorders such as sleep apnea or narcolepsy may be advised by physicians to undergo sleep studies in a laboratory, which consists of a single, highly studied night. PSG is the clinical standard for diagnosing sleep disorders by sensing sleep and its various stages. An entire night's data (comprising of a minimum of 6 hours of sleep time) is analyzed for 30 second intervals from a combination of signals derived from sensors such as: an electroencephalogram (EEG, used to sense brain waves), electrooculography (EOG, used to track eye movements, electromyography (EMG, used to capture electrical activity produced by skeletal muscles), pulse oximeters, microphones placed under the chin(to monitor snoring and teeth grinding events), RIP (respiratory inductance plethysmography belt to monitor diaphragmatic breathing) and a nasal canula (to measure air flow through the nostrils) [72]. The various stages of sleep are classified based on specific time domain patterns as well as frequency characteristics in the EEG signal. Scoring is manually performed by a trained sleep technician by visually identifying signal characteristics in 30 second intervals called epochs. These stages are classified based on characteristic patterns of brain waves which can be recorded through an EEG and monitoring eye movement using an EOG which form a part of the montage in a Polysomnograph. Although the pattern of these brain waves and onset of eye movements are the most definitive indicators of sleep stages, the other sensors provide valuable information to diagnose sleep related disorders such as grinding of teeth, sleep apnea and restless leg syndrome.

Sleep data from an overnight sleep study is usually represented as a hypnogram. A typical hypnogram starts with the patient being awake and cycles through the 5 sleep stages. PSG provides fine-grained sleep quality assessment. PSG is considered a highly obtrusive sleep sensing system due to its expense, impracticality for home-based use, and comfort-level for the patient, and thus its application is limited to diagnosing sleep-related disorders in clinical settings for a short duration.

### *2.2.2 Actigraphy*

However, many people with poor sleep quality have more long-term sleep issues such as insomnia or delayed sleep phase syndrome that require longer term monitoring to diagnose, monitor, and treat. Therefore, sensing sleep and tracking sleep patterns over long periods of time can help bring awareness and play a critical role in the treatment of sleep related disorders. For long-term sleep tracking, actigraphy has been established in the medical community as an acceptable tool to continuously track sleep in a non-clinical or home setting. The term Actigraphy refers to methods using miniaturized wrist-watch-like devices with accelerometers embedded in them to monitor and collect data generated by body movements over several weeks [9]. Actigraphy has been used to study sleep/wake patterns for over 20 years. The advantage of actigraphy over traditional polysomnography (PSG) is that actigraphy can conveniently record continuously for 24-hours a day for days, weeks or even longer [19]. Several review papers have concluded that wrist actigraphy can usefully approximate sleep versus wake state during 24 hours and have noted that actigraphy has been used for monitoring insomnia, circadian sleep/wake disturbances, and periodic limb movement disorder. In healthy adults, actigraphy was valid for assessing sleep duration and sleep/wake activity, but less reliable for more specific measures such as sleep offset or sleep efficiency. When compared to PSG, actigraphy was found to be valid and reliable for detecting sleep in normal, healthy adult populations but less reliable for detecting sleep as sleep became more disturbed [111]

### *2.2.3 Consumer Sleep Tracking Devices*

In recent years, numerous consumer electronic fitness trackers leverage inbuilt accelerometers to provide actigraphy-like sleep monitoring. The raw accelerometer data is continuously sampled to detect coarse body movements such as tossing and turning during sleep and marks these as arousal events. The widespread availability and adoption of these devices has led to an increased interest in sleep tracking for non-clinical purposes [79]. Dedicated consumer sleep sensing devices are available in several form factors and utilize a wide variety of sensors for sensing sleep related physiological parameters. Sleep sensing devices have the potential to help people learn and monitor their sleep patterns over long periods of time outside a clinical setting. This awareness can be attributed to the fact that consumer sleep sensors offer users immediate feedback, in contrast to clinical actigraphy devices that generate data that can be only be viewed by clinicians. A formal review of the consumer sleep sensing landscape was done by Ping-Ru et.al [75] along with their effect on clinical sleep medicine, and various social, legal, and ethical issues. In this section, I have briefly outlined the different types of consumer sleep sensing devices in-order to underscore some of their hardware and algorithmic limitations.

#### *Wearable Sleep Trackers*

Most wearable sleep sensing devices are mostly available in a wrist-watch like form-factor similar to clinical actigraphy devices. These devices leverage data from several sensors like the Inertial Measurement Unit (IMU) sensors for detecting motion, photo-diodes for pulseplethysmogrpahy or PPT or an EKG for eclectrocardiac acvitiy and sometimes even a galvanic skin response (GSR) sensor to measure electrodermal activity. Together, these sensors measure coarse body movements, heart rate, abnormal heart beats, and stress. As the body cycles through different stages of sleep, these sensors measure the changes associated with these physiological signals corresponding with the cycles to estimate sleep stages and assess sleep quality. A few popular dedicated wrist worn sleep trackers are the Withings Move [14], Polar M430 [8], and the Fitbit Versa [3]. In addition to dedicated devices, smart-watches like the Apple watch allow third-part applications to track users' sleep [1], while the Samsung Galaxy smartwatch [11] comes with built-in sleep tracking software.

### *Accuracy of Wearable Sleep Trackers*

These devices are not clinically validated, and their accuracy compared to the gold standard are not made public by manufacturers. For example, Montgomery-Downs et al. compared the accuracy of Fitbit and ActiWatch against a PSG study [97]. The authors found that Fitbit and ActiWatch differed significantly on recorded total sleep time, both between each other and compared to PSG. While, newer versions of these devices with improved sensing and algorithmic capabilities continue to flood the marketplace as smartwatches steadily gain popularity, the accuracy of these devices still remains an issue [52, 44]. The most common issue being the IMU sensors' limitation at discerning between laying still but awake in bed and sleeping. Sleep stage assessment from physiological biomarkers of sleep can never be as precise as estimating sleep stages from brain waves (as done during a PSG study).

### *Hardware Limitations of Wearable Sleep Trackers*

In addition to wrist-worn sleep trackers, commercial sleep sensing devices are available in a few other form factors such as the Oura ring: a sleep tracker that can be worn on the finger [6], the Philips Smart-Sleep deep sleep headband [7] and the Zeo headband [15] that come with built in 3-electrode EEG sensors and an EOG sensor. The sensors in the headband and the eye mask offer better sensing capabilities since they directly measure brain waves and eye movement associated with the different sleep stage. In spite of their improved accuracy an issue that plagues almost all wearable sleep sensing devices is their intrusiveness. Most users find it cumbersome to wear devices on their body while they sleep [29]. Liu et al. [91] conducted an investigation of the usability and acceptability of commercial sleep sensing devices and reported that discomfort, battery life, and inability for users to modify their data are major limitations in current sleep sensing devices. Studies have also found that the lack of definition of sleep metrics, limitations in underlying data collection and processing mechanisms reduce users' trust in sleep sensing devices [88].

### *2.2.4 Mattress-based Sleep Trackers*

In order to overcome some of the limitations of wearable devices, several mattress-based sleep sensors have been developed such as the Beddit [2] and the Withings Aura [13]. These devices contain a thin flexible,sensor placed under the sheet on top of the mattress. They track sleep parameters such as duration and efficiency, heart rate, respiration, temperature, movement, snoring, and also environmental variables that affect sleep such as room temperature, and humidity. Since these sensors are pressure-based, there's no need for the user to wear it on their body or recharge it. The Sense is a spherical device that can be placed on the bedside and tracks sleep by pairing with the 'Sleep Pill', a coin shaped device that attaches to a pillow. The Pill features a 6-axis accelerometer and gyroscope which detects body movement during sleep. The Sense also has a range of environmental sensors to monitor the amount of light, noise and also the air quality [12]. While these device eliminate the need for the user to wear a device on their body overnight, they are still restricted by range and accuracy. Users can fall out the range of these sensors by rolling over to the far side of the bed over the course of the night. Also, since these devices are pressure based, the sensitivity of the devices largely affects their accuracy.

### *2.2.5 Smart-phone-based Sleep Sensing Applications*

In order to facilitate non-contact and unobtrusive sensing, several smart-phone apps have been developed to assess sleep quality and track sleep sleep over long periods of time. Smartphones have become ubiquitous and are generally in the proximity of users even at night. By placing a smart phone beside the user on the bed, the inbuilt sensors in the smartphone can monitor environmental variables that affect sleep, coarse body movements, as well as sounds from sleep related breathing disorders. Hao et al. proposed iSleep [57], which leverages a smartphone's built-in microphone to unobtrusively measure sound caused by body movement, cough, and snoring. Gu et al. proposed Sleep Hunter [53], a system that uses the smartphone's microphone, accelerometer, light sensor, etc.

to capture both environmental disturbances (light, noise etc.) and human physiological reactions (like movements, cough, snore etc.) to model different stages of sleep (e.g., REM, deep, and light sleep). In Toss 'N' Turn, researchers investigated the accuracy of sleep sensing using data from several in-built sensors found in smart-phones to measure light-intensity, acceleration, sound amplitude,list of running apps, phone battery state, etc., and found it was possible to predict aspects of sleep quality to between 81-83% accuracy [96]. Nandakumar et.al have developed a smartphone based sleep apnea detection system that leverages active SONAR emitted by the phone's speaker to detect apena events (short periods of interrupted breathing) during sleep [99]. A study conducted in 2016 showed that there were at least 51 unique sleep related smart phone apps available in both iOS and Google Play stores. A majority of these sleep apps (65%) reported on sleep structure, including duration, time awake, and time in light/deep sleep and a few reported on REM as well [101]. While enabling unobtrusive sleep monitoring, smartphone based applications do not accurately capture sleep related bio-markers due to their sensor limitations. Studies that compared smartphone based sleep apps with that a PSG found that there was no correlation between sleep efficiency scores, light sleep percentage, deep sleep percentage or sleep latency estimates generated by the smartphone app and the PSG [23].

### *2.2.6 RADAR Based Sleep Sensing*

The concept of contactless vital signal monitoring using microwave signals has been explored since the 1970s [90]. Since then a lot of research has been done to improve the performance of the system both in analog circuit design of high frequency carrier signal as well as signal processing of basebanded signal. Different choices of radio frequencies have been explored starting from 1150 MHz ultra-wide band to detect vitals through earthquake rubble and concrete [28] all the way up to Ka band [120] to improve detection sensitivity. Different receiver architectures [84] and techniques to compensate for phase noise [82] have been proposed and tested. Over the years, RADAR modules have been reduced from bulky military grade systems mounted on a tripod to a relatively small BiCMOS chips suitable for integration in portable electronic devices [45]. These fully integrated devices with much smaller form-factors have opened avenues for the use of electromagnetic signals for near-field sensing. Using electromagnetic signals for vital signal sensing has the added benefit of being able to sense vital signals through multiple layers of clothing or textile such as blankets. The reflectivity of the human body to electromagnetic signals allows minute thoracic movements from breathing and heart beat as well as coarse body movements to cause changes in the propagation path of the signals. This extends the range of sensing considerably compared to the smartphone based sleep sensing applications. Features extracted from these sleep related physiological biomarkers can be used to estimate sleep parameters. The S+ [10] is a commercial device that measures breathing and body movements throughout the night through a pulsed radar wireless signal. It also analyzes light, noise and temperature levels in the bedroom. The device then computes a sleep score and provides feedback regarding sleep patterns on a phone-app. While the S+ provides the least intrusive solution to sleep monitoring, it's accuracy is still not comparable to the clinical sleep monitoring systems since it doesn't measure the third sleep related physiological signal: heart beats. Resting heart rate as well as changes in heart rate variability through out the night can provide key information regarding sleep quality and the transition between different sleep stages.

### *Overview of Sleep Sensing*

Given the popularity of consumer sleep tracking devices with both the general population and with people suffering from sleep disorders, studies have tried to provided an overview of the different categories of consumer sleep tracking devices [75], discussed their advantages and disadvantages, and explored how clinicians may incorporate these devices in their interactions with patients [112].

In Figure 2.1, I have provided an overview of the different sensing solutions for sleep categorized by intrusiveness. Starting from the highly intrusive, medical gold standard for sleep sensing used for diagnostic purposes in the right to the least intrusive daily tracking solutions on the left. While the clinical gold standard has the highest accuracy, most of the commercial devices available for daily tracker have much lesser, similar accuracy since they share similar sensing modalities. In chapter 6 I have explained how my approach of using a 5GHz continuous wave RADAR system can be effectively used to estimate sleep quality by measuring three sleep related biomarkers such as breathing, heart beat and body movements in an unobtrusive manner.

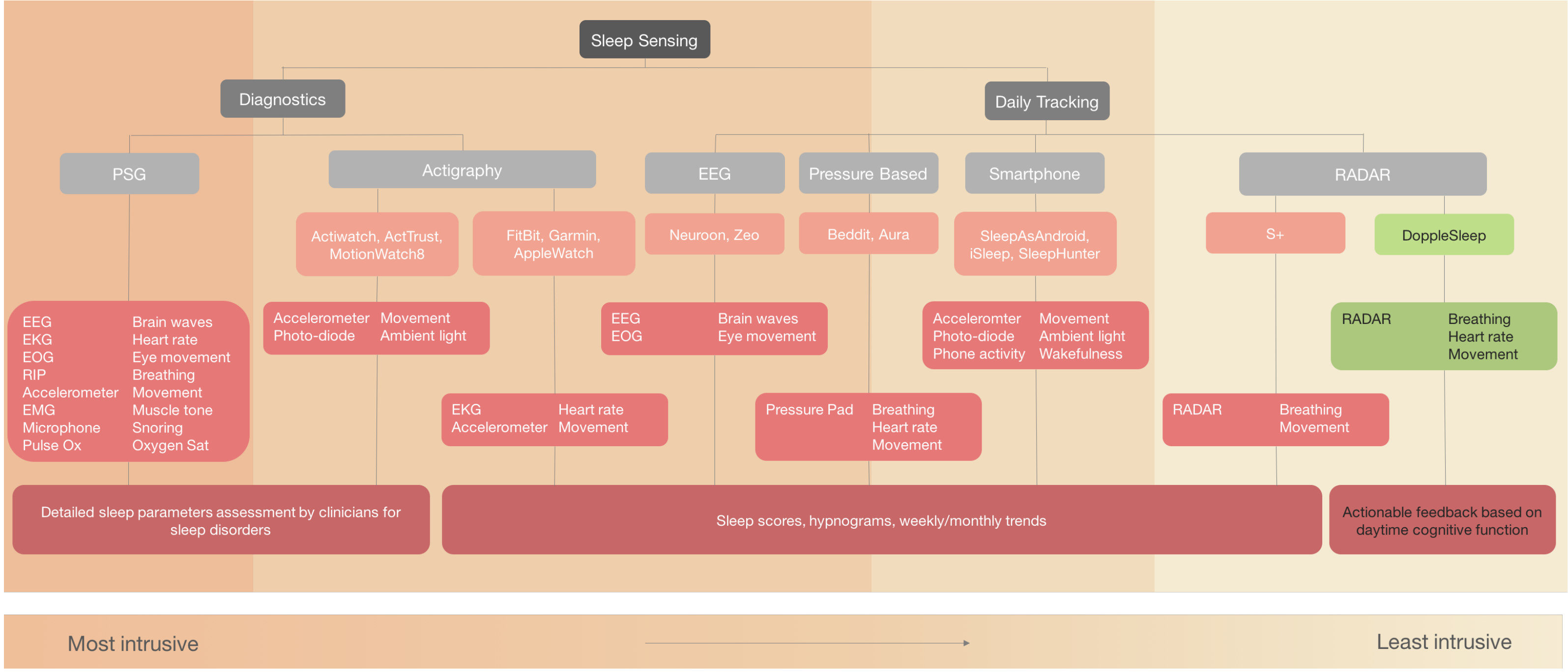


Figure 2.1: An overview of the different methods and sensors used for sleep sensing starting from the most intrusive on the left to the progressively less intrusive towards the right.

### *2.3 Sleep Metrics*

An important component of measuring physiological signals related to sleep, is quantifying them in such a way that it adequately represents the quality of sleep. However, sleep cannot be quantified in a simple, straightforward manner and represented using a single value like other health measures. While the general population uses standardized terms such as BMI or h1bac to assess physical health, quantifying sleep does not have an equivalent standard measure and qualifying it effects can vary drastically from person to person. Buysse et al discuss in [26] the lack of a clinical definition of sleep health. The author highlights that Principles and Practice of Sleep Medicine [77] mentions "sleep health" twice, but does not define it.

While the medical community uses several metrics to define sleep quality and interpret results from a PSG study, they are largely been focused on the identification, and treatment of sleep problems such as sleep apnea, periodic limb movement disorder, parasomnias and narcolepsy. The data captured from the PSG study is used to assess the following measures: sleep efficiency (the ratio of total sleep time to time spent in bed), sleep latency (the duration from bedtime to the onset of sleep), arousal index (number of awakenings after sleep onset), number of minutes spent in the various sleep stages, and other metrics to assess sleep apnea [72]. While these measures help physicians diagnose sleep-related disorders, on their own, however, they are insufficient to assess sleep quality for people wanting to track their sleep long-term or to diagnose behavioral sleep disorders, such as insomnia. Especially, when the same metrics used in clinical settings and interpreted by medical professionals are used to provide feedback to lay users who track their sleep over long terms to determine overall sleep health.

### *2.4 User Perception of Sleep Feedback*

In the broader field of self tracking technologies, there have been studies of how people make sense of their self-tracked health data and why people abandon self-tracking when they struggle to make sense of said data [48, 68]. Previous work has examined how to represent and visualize data such that it is persuasive and offers insights that can lead to behavior-change [35]. Uncertainty is a point of frustration for users of physical activity inference technologies. Users of these technologies have to cope with activity inference and measurements that are prone to error. Consolvo et al.

identified that users react negatively when fitness trackers incorrectly infer a particular physical activity and consequently, do not give users credit for said activity [35]. Kay et al. found that there is a disconnect between a users' perception of their weight, the precision capabilities of their scale, and clinical relevance of weight deviations [68]. Kay et al. further found that an accurate understanding of weight fluctuation is associated with greater trust in the scale itself. Work by Yang et al. [122] examined how self-trackers view the inaccuracy of sensor-driven step count inference and the process in which self-trackers engage to assess the accuracy (or lack thereof) of their fitness devices.One such way to improve data representation is to clearly convey its uncertainty [66]. These studies demonstrate users care about the accuracy of sensor-driven tracking, taking accuracy into account when they assess their data. While the insights from these studies can be useful and applied to sleep sensing technologies as well, there are some unique challenges that come with interpreting sleep data especially because sleep is an unconscious activity. It is important to understand what aspects of sleep sensing and feedback either facilitate or potentially undermine people's ability to understand their sleep and achieve good sleep health and if current sleep sensor technology designs are in line with evidence-based methods of understanding and promoting good sleep health.

## *2.5 Sleep Hygiene*

Sleep quality can be affected by several variables. They can be environmental, physiological or behavioral. To better understand and improve sleep health, we need to take into the account the effect of these variables. For example, to improve sleep quality, sleep experts and literature review suggest users assess sleep habits and adopt modifiable behaviors [63]. Modifiable behaviors are behaviors that people have control to act on, which include examples such as: (1) keeping one's bedroom cool and dark; (2) maintaining a regular bedtime and wake time every day, even on weekends; and (3) avoiding large late-night meals. A second and related concept is sleep hygiene, which refers to behaviors, habits, and environmental factors that can be adjusted to promote good sleep quality [115]. Examples of sleep hygiene include avoiding caffeine later in the day, exercising regularly, and establishing a relaxing bedtime routine [63, 115]. Addressing modifiable behaviors and sleep hygiene are the first two methods sleep clinicians use when patients complain about poor sleep quality.

## *2.6 Sleep Tracking*

In the HCI and UbiComp communities, researchers have developed a number of novel computing-based interventions for sleep. In 2011, Choe et al. conducted a literature review and formative study to examine design opportunities for sleep from an HCI perspective [29]. The authors identified people's strong interest in lowering the barriers to track their sleep and the factors that affect their sleep. The authors also stressed the importance of supporting long-term sleep tracking to identify trends to help people create personalized sleep goals. Building upon these opportunities, researchers have explored varying ways of capturing and providing feedback on aspects related to sleep health. Cognitive Rhythms [17] implemented an in-the-wild study to measure how day-time alertness varies between people of different chronotypes. SleepTight [30] is a mobile application that lowers the barriers of manually tracking sleep and helped users make sense of behavioral factors that could be affecting their sleep. ShutEye [21] is a peripheral display on a smartphone's active wallpaper that provides timely guidance on when it is best to engage in activities that could impact sleep, such as consuming caffeine or exercising. Lullaby provides comprehensive information of users' sleep environment. This information allows users to learn about environmental factors that may affect sleep [65]. Lullaby captures environmental factors (e.g., temperature, light, audio, and motion) in relation to sleep data captured from a Fitbit. Although these technology designs provide valuable information to help users understand their sleep, they primarily focus on sharing collected information to people and make recommendations based on population-level data.

There still remains a gap in understanding the extent to which behavioral changes to the variables affecting a person's sleep can have a direct correlation with well-being and cognitive abilities.

## *2.7 Self-Experimentation for Sleep*

Connected to the design of self-tracking technologies is the study of how people make sense of their self-tracked health data and why people abandon self-tracking when they struggle to make sense of said data [48]. To help people make sense of their data, research has explored guiding users through a process of self-experimentation. Karkar et al. have provided an overview and framework on self-experimentation [64], and others have worked to develop self-experimentation tools specifically for sleep. In SleepCoacher, Daskalova et al. [42] presented a framework for effective sleep-based self-experiments. SleepCoacher automates the cycle of single-case experiments by collecting raw mobile sensor data and generating personalized, data-driven sleep recommendations from a collection of template recommendations created with input from clinicians. The system guides users through iterative short experiments to test the effect of recommendations on their sleep. However, the study was limited in that the effects of poor quality of sleep was measured only using a subjective self-report rating.

In Self-Experimentation for Behavior Change, Lee et al. [80] report on the iterative design of two complementary support strategies for helping users create their own personalized behavior-change plans via self-experimentation: One emphasized the use of interactive instructional materials and the other additionally introduced context-aware computing to enable user creation of "just in time" home-based interventions. Lee et al.'s work investigated how to help people develop a behavior change protocol using habits developed using triggers and SMART (Specific, Measurable, Actionable, Realistic, and Timely) goals. We incorporated findings from this work in our study design and chose appropriate measurable, realistic, and actionable constructs such as day-time alertness and working memory as the outcomes to measure.

Finally, Daskalova et. al. [40] conducted a study on self-experimentation with two cohorts of students, where a total of 34 students performed a self-experiment of their choice. The authors observed how novices run self-experiments and provide guidelines on self-experiment design in terms of choosing the right variables, the length of the study, affects the way people run self-experiments and analyze their data. Further, Daskalova et al. [41] later found that social comparison inspired cohort-based systems for health recommendations towards behaviour change and increased awareness about sleep habits. The findings from these studies served as building blocks for our study design, as our ultimate goal was to provide a method and basis for people to self-experiment on factors of sleep that result in measurable impacts on their life, such as mood, alertness, and cognitive functioning.

### *Ru-SATED measure*

A new framework by Buysse et al. [26] proposed to define sleep health as "a multidimensional pattern of sleep-wakefulness, adapted to individual, social, and environmental demands, that promotes physical and mental well-being. Good sleep health is characterized by subjective satisfaction, appropriate timing, adequate duration, high efficiency, and sustained alertness during waking hours." Based on this definition, Ru-SATED is a framework that uses six dimensions to measure holistic sleep quality as follows:

- Regularity: the consistency of bedtime and wake time
- Satisfaction: the subjective assessment of "good" or "poor" sleep
- Alertness: the ability to maintain attentive wakefulness
- Timing: the placement of sleep within the 24-hour day
- Efficiency: the ease of falling asleep and returning to sleep
- Duration: the total amount of sleep obtained per 24 hours

An associated scale with this framework has each measure is rated on a scale from 0-2 with the total score possible score of 12. The Ru-SATED scale has been validated against the Pittsburgh Sleep Quality Index (PSQI) [27] and was found to be reliable for the assessment of sleep health [24]. This framework forms the base from which meaningful correlations between sleep quality and daytime cognitive function can be generated.

## *2.8 Guidelines from Personal Informatics*

Personal informatics covers a class of systems that help people collect and reflect on personal information. In the HCI community several studies have been conducted to characterize the different aspect of self-tracking for health and well-being.

### *2.8.1 Personal Informatics Models*

Some of the formative work in personal informatics was done by Li et al [85], who made a comprehensive list of problems that users' experience with self tracking systems. They derived a stage-based model of personal informatics systems composed of five stages: preparation, collection, integration, reflection, and action. They identified four essential properties and barriers that cascade to later stages. Rooksby et. al. [108] characterised the use of activity trackers through the term ‘lived informatics’, featuring two types of tracking: goal driven tracking and documentary tracking. They identified that tracking information is often used and interpreted with reference to daily or short term goals and decision making. Epstien et al developed a model to characterize the self-tracking processes with a focus on the integration of self-tracking into everyday life by people with varying goals. They proposed a lived informatics model consisting of the decision to track, selection of a tool, tool usage for data collection, integration and reflection on the data as well and lapsing and resuming tracking [49]. In a following paper, Epstien et al also studied why people stop using personal tracking tools over time and provide design guidelines to support people when they lapse in tracking [47].

### *2.8.2 Challenges in Long Term Tracking*

Li et. al. [86], explored users’ need for self-reflection from their data and identified two phases of reflection: Discovery and Maintenance. They then identified features that should be supported in personal informatics tools to support self-reflection. Meyer et. al. [95] specifically observed the characteristics and challenges of long-term self-tracking for health. They found that changes in users’ goals and their commitment to tracking affected adherence. They also identified that the type, amount and the quality of data were determined by the user’s tracking routine and not by potential future needs. The highlighted key challenges for designing applications for long-term tracking such as matching the required effort with anticipated benefits.

## ***2.9 Summary***

This large body of work forms the foundation upon which I have attempted to build my research. Since sleep tracking spans several domains such as biology, medicine, engineering and HCI, I have attempted take a holistic view and make contributions across the different domains. I positioned my research contribution in creating a novel sensing solution within the context of the hardware and accuracy limitations of the current commercial devices. By surveying contributions in the HCI community around sleep tracking, and sleep self-experimentation, I built upon their findings with a specific contribution in providing sleep feedback that is contextualized with day-time cognitive function. In the following chapters I have discussed how sleep feedback can be provided in a meaningful way to promote sleep health and how a novel sleep sensing device can enable unobtrusive, long term sleep tracking.

# Chapter 3

# UNDERSTANDING USER PERCEPTION OF SLEEP DATA

## *3.1 Introduction*

In this Chapter, I describe the study we[1] conducted to understand the gaps between user perception of commercial sleep sensors and expert views on good sleep practices. This study forms the foundation for the following contributions in unobtrusive sensing as well in providing actionable feedback by answering the first research question:

- RQ1: How do users perceive the feedback they receive from current commercially available sleep sensing devices?

To understand people's perspectives on sleep sensing devices, we surveyed 87, interviewed 12 people who currently use or have previously used sleep sensors and conducted an in-depth qualitative analysis of 6986 reviews of the most popular commercial sleep sensing technologies. To understand expert opinion on good sleep practices, we interviewed 5 sleep medical experts and conducted review of medical literature. The findings of the study highlight aspects of sleep sensors that act as facilitators to good sleep health and those that act as barriers. Based on these finding we provide design guidelines to help remove barriers to promoting good sleep health.

## *3.2 Motivation*

Commercial sleep sensing technology for use at home is a growing industry [4]. These technologies have the potential to overcome the limitations of PSG studies while providing long-term, low-cost, and accurate representations of people's daily sleep patterns in their natural and comfortable home environment. The popularity of these commercial sleep sensors is promising in that they indicate that people have an interest in understanding and obtaining good sleep health. However, literature

---

[1] This study was conducted in collaboration with the following co-authors: Sang-Wha Sien, Shwetak N. Patel, Julie A. Kientz and Laura R. Pina.

has not examined whether commercial devices effectively sense sleep quality and provide people with meaningful feedback. Specifically we wanted to answer the research questions:

- How are people currently using commercially available sleep sensors and making sense of feedback they provide?
- What aspects of sleep sensing and feedback either facilitate or potentially undermine people's ability to understand their sleep and achieve good sleep health?
- What aspects of current sleep sensor technology designs are in line with evidence-based methods of understanding and promoting good sleep health?

To answer these questions, we collected a data-set consisting of interviews with 5 sleep experts, surveys with 87 and interviews with 12 people that have used sleep sensing devices, and 6986 consumer product reviews from the most widely used commercial sleep sensing devices. We focused on sleep sensing technologies that use physiological sensing, such as body movement, breathing rate, or heart rate to estimate sleep quality and excluded manual, self-reported sleep tracking methods such as sleep diaries. We find that:

- Self-trackers using sleep sensing technologies often develop broken mental models about what commercial sleep sensors are able to actually sense, how they work, and are frustrated with the lack of algorithmic transparency in sleep sensing technologies.
- Self-trackers find it distracting when feedback emphasizes unconscious aspects of sleep, such as time in sleep stages, over aspects of their sleep they have the ability to control and improve.
- Self-trackers can better understand and improve their overall sleep habits when feedback from sleep sensors focuses on duration, timing, and making connections to modifiable behaviors and sleep hygiene.

Our findings examine the state of sleep sensing feedback from the perspective of users' needs and sleep experts. From our results, we derive design recommendations that consider users' needs and connect them to evidence-based strategies for improving sleep quality. A set of these recommendations provides new avenues to improve sleep sensing.

### *3.3 Data Collection*

To understand the state of sleep research and the needs of people using sleep sensing devices to track their sleep, we 1) interviewed sleep experts and reviewed the literature on sleep research, 2) analyzed consumer product reviews of sleep sensing devices, 3) deployed an online survey, and 4) interviewed a subset of survey respondents. In this section, we describe our process and analysis methods.

#### *3.3.1 Interviews with Sleep Experts*

To gain an understanding of the factors contributing to sleep health, we conducted a literature review of sleep research and interviewed five experts in the field of sleep medicine (E1- E5). E1 is a Neurology professor and board certified sleep specialist. E2 is a professor in Psychiatry and Behavioral Science, co-director of a sleep research center, and editor of a major sleep research journal. E3 is a sleep researcher in a department of Family and Nursing. E4 is a professor in a department of Family and Child Nursing and focuses on pediatric sleep. Finally, E5 is a pediatric psychologist and sleep researcher. Experts were familiar with commercial sleep sensors and the feedback they provide. These interviews helped us understand experts' perspectives on how sleep sensing technologies address sleep health needs and the practices the experts establish with patients who use sleep sensing technologies to track their sleep. During the interview, experts were asked to comment on feedback examples and discuss how they use patient-generated sleep sensing data. We analyzed the sleep expert interviews with support from the sleep literature to identify themes focused on maintaining and improving sleep.

#### *3.3.2 Reviews of Sleep Sensing Products*

We collected and analyzed product reviews from the most widely-used commercially available sleep sensing technologies to gather a user perspectives on sleep sensing 'feedback. We gathered reviews from three sources: Amazon.com, iTunes Store, and Google Play Store. Our inclusion criteria consisted of: 1) smartphone apps using phone sensors (e.g., accelerometer and/or microphone), 2) dedicated sleep sensing devices, or 3) fitness trackers which also sense sleep. For smartphone apps, we analyzed reviews from the 4 highest-rated apps from the iTunes Store and the 5 highest rated apps from Google Play. We selected reviews in decreasing order of word count (e.g., longest reviews first), stopping once we felt we reached data saturation. For iTunes reviews, we reached data saturation at 280 word count, analyzing 475 reviews out of a total of 2000 possible reviews. For Google Play reviews, we reached data saturation at 500 word count, analyzing 377 out of a total of 14581 possible reviews. Combining both sources, we analyzed 852 app reviews. From Amazon.com, we collected reviews from dedicated sleep sensing devices. These are sensors that are placed under the mattress (e.g., Beddit, Withings Aura), clipped on the sleeper's pillow (e.g., Sense with Sleep Pill), or placed on the nightstand (e.g., S+). We analyzed all 683 reviews for these five dedicated sleep devices. Also from Amazon.com, we collected reviews from the top four suggested wearable fitness trackers with sleep sensing functionality: Fitbit One, Fitbit HR, Jawbone Up3, and Misfit Shine. We only included fitness tracker reviews containing the word 'sleep'. This led to 3234 Fitbit One, 4298 Fitbit HR, 893 Jawbone Up3, and 78 Misfit Shine reviews to analyze. Similar to our data saturation process for the smartphone app reviews, we read reviews in decreasing order of word count, analyzing data until we felt we reached data saturation. These reviews tended to be longer than the smartphone app reviews. The three authors coding this dataset reached saturation at different word counts for some of the devices. In total, we analyzed 2113 Fitbit One, 2452 Fitbit HR, 808 Jawbone Up3, and all 78 Misfit Shine reviews, totaling 5451 fitness tracker reviews. Combining all review datasets, we analyzed a total of 6968 reviews

### *3.3.3 Online Survey*

The themes identified from the expert interview data and the review dataset informed the list of questions to survey selftrackers using sleep tracking technologies. The 29-question survey focused on: 1) reasons why people track their sleep, 2) which sleep sensing devices people use and why those devices, 3) the type of information people wanted to collect, 4) how people make sense of the feedback from sleep sensing technology, and 5) how people connect data to their sleep quality. Questions were a mix of open-ended, Likert, and multiple choice. We recruited by posting on social networking sites, online message forums, and through a sleep blog. To incentivize participation, respondents were entered into a drawing to win one of five $20 USD Amazon gift cards. We gathered a total of 87 responses.

### *3.3.4 Semi-Structured Interviews*

Survey respondents had the option to consent to be contacted for an in-depth follow-up interview. We contacted all 46 respondents that consented. We interviewed the 12 which replied to our request. We conducted interviews over the phone or in person. Interviews lasted between 16 to 30 minutes. With consent from participants, we recorded and transcribed interviews. Interview questions were based on respondents' survey answers and were intended to triangulate and add depth to our findings from the survey, app reviews, sleep literature, and interviews with experts. Five interviewees had been diagnosed with a sleep disorder and three had stopped tracking. We compensated interview participants with a $25 USD Amazon gift card.

## ***3.4 Analysis***

Our analysis consisted of an iterative affinity diagramming process with 6 steps to analyze our triangulated dataset [4]. In Step 1, we analyzed expert interview data and the literature. We identified 7 themes focusing on sleep hygiene, modifiable behaviors, experts' perspectives on how sleep sensing feedback can help their patients address sleep concerns, and how patients and physicians use feedback provided by sleep sensing devices. In Step 2, we analyzed the product review dataset, which generated 64 themes. In Step 3, we created our survey based on the themes generated from the two previous steps. In Step 4, we analyzed the survey data and merged it with the themes identified from the product review dataset (i.e., Step 2). In Step 5, we applied the 7 themes from the expert data to the themes generated from the survey and review dataset, but kept themes reflecting user practices and challenges. This step trimmed our themes 64 to 30. Based on these themes, we created our interview protocol to gather deeper insights. Finally, in Step 6, we integrated the interview data to identify higher level themes presented in the results. For every step of the analysis that required affinity analysis, the data was split between three authors. Each author analyzed their subset of the dataset. We then came together to merge, discuss, and iterate on themes.

## ***3.5 Sleep Sensors as Facilitators***

Sleep sensing feedback provides awareness, motivates users to prioritize sleep, helps improve sleep habits, and helps people with sleeping disorders collaborate with their physicians to better manage their condition. In this section, we discuss the strengths of sleep sensing feedback and opportunities to improve.

### *3.5.1 Promoting Awareness about Sleep Health*

Inadequate sleep is the most common sleep issue in the United States [56]. Experts in our study agreed, explaining that sleep is a low priority for people: “I don’t think that people pay much attention to sleep until they have a problem” (E3). All experts in our study said sleep sensing technologies can create awareness of the importance of sleep in typically healthy people: “I don’t think sleep hits their radar unless someone actually shows them saying look, you are not getting a lot of sleep” (E4). Just from the sheer availability of consumer sleep sensing technologies, people have started to utilize these emerging technologies to learn more about their sleep habits. 83.9%(73/87) of our survey respondents and a majority of online reviewers considered themselves healthy, were very interested in understanding their sleep, and discussed the benefits of having access to information about their sleep. This provided users with information they were previously unaware of.

### *3.5.2 Facilitating Adoption of Healthy Sleep Habits*

In addition to increasing duration, sleep quality plays an equally vital role in health and well-being. To improve sleep quality, one must address modifiable behaviors. Sensors that capture environmental factors such as acoustical noise, room temperature, and ambient light help users identify potential environmental factors that may be impacting their sleep. Compared to other health conditions, sleep quality is highly subjective. The number of hours, modifiable behaviors, and changes in sleep hygiene to improve sleep quality may vary for every person.

Although people believe it is important to prioritize sleep and taking proactive steps to address sleep quality, they also have their own beliefs and metrics on sleep. Personal beliefs can be carefully examined by incorporating a self-assessment framework. The SATED framework [26] can be used to identify the quality of a person’s sleep and personalize what adequate sleep means to a specific user.

### *3.5.3 Managing Sleep Disorders or Chronic Conditions that Affect Sleep*

Feedback from sleep sensing also has the potential to manage sleep disorders. This data can help experts work with patients to identify and manage patients' sleep conditions. Sleep sensors can improve assessment and screening. Because PSG studies take place in a clinic, they do not represent a patient's natural sleep environment. They are therefore not well-suited to study non-physiological disruptors of sleep. Sleep sensing feedback can help patients determine the effectiveness of a treatment for a particular sleep disorder, such as using a CPAP (Continuous Positive Airways Pressure) machine for sleep apnea. Experts stated that patients struggle to adhere without longitudinal data on the effects of the treatment: "Is the treatment working? Something beyond their subjective sense of whether or not they're better, but some objective data to show that their sleep quality is better" (E1). Connecting sleep with treatment effectiveness is crucial, especially since over time, patients' motivation to adhere to treatment decreases: "… they forget what it was like before, they get uncertain as to whether or not they're better. The device might be able to increase that certainty to motivate ongoing compliance with treatment" (E1).

## ***3.6 Sleep Sensors as Barriers***

Although sleep sensing devices provide useful and objective feedback that is beneficial to users, our analysis identified areas of improvements and opportunities incorporate evidence-based strategies to sleep sensing feedback. Feedback from the sleep sensors tends emphasize estimating the number of hours users spent in various sleep stages and assessing sleep quality using computed, single-point measures such as Sleep Efficiency or Sleep Score. Current feedback tends to focus on these sleep stages. For neurotypical people, a breakdown of the time spent in the different sleep stages is not helpful feedback for improving sleep quality. The focus on sleep stages leads to users developing inaccurate mental models of how current sleep sensors work and what it means to get good quality sleep. Feedback on sleep stages distracts users from focusing on adapting modifiable behaviors to improve sleep hygiene, which really could have a positive impact on their health.

### *3.6.1 Inconsistency in Sleep Quality Inference*

Experts noted that there is a trait variability to sleep and some measures are very specific to each person, "Some people can … have poor quality sleep and not really feel many ill effects from it. Other people can have just minor decrements in those sleep factors that have a pretty big impact on sleep quality (E1)." To provide feedback to users about their sleep, commercial sleep sensors often focus on determining objective measures such as sleep efficiency, sleep latency, and the different stages of sleep. Commercial sensors tend to focus less on subjective measures. However, the focus on objective measures led many users to have a broken mental model of what sensors can infer and what information is useful to address sleep concerns. R2066 (Fitbit One) says, "The lack of explanation as to the formula/algorithm that lead to the results are very maddening. What does the 94% effective sleep rating actually mean?"

### *3.6.2 Mismatch between score and user perception*

In our survey, 50% of our respondents agreed or strongly agreed that their sleep score or sleep efficiency was related to their sleep quality (Figure 3.1). However, these scores may not necessarily provide users an accurate picture of their sleep quality because sleep efficiency scores vary on hardware specification and sensing sensitivities. "The problem is that it is nowhere near sensitive enough on normal and way too sensitive on sensitive setting. I had a restless night the night I had it on normal, waking multiple times, and it recorded 15 minutes of restless sleep and no wake times. The next night, on sensitive, I had much better sleep, and it recorded only 3 1/4 hrs. of sleep and the rest waking or activity!" (R10, Fitbit One).

### *3.6.3 Placing Undue Emphasis on Sleep Stages*

Reviewers of sensing devices place high value in sleep sensing devices that can infer sleep quality based on sleep stages. One review said: "The only thing it doesn't really do that my mom's Jawbone does, is that it doesn't tell me about my sleep cycles (stage 1/stage 2/REM etc.) it just tells me when I'm asleep, awake, and 'restless'" (R784, Fitbit Charge). Survey respondents reflected the same perspective, considering sleep stages to be representative of sleep quality (see Figure 3.1). More than 60% agreed or strongly agreed that time spent in specific stages such as REM, deep was related to sleep quality. I8 said, “The Jawbone … gave me light sleep and the deep sleep separation and, of course, the awake states. Since I didn’t have many awake states during the night, which was good, I only pretty much had the amount of hours that I was having light sleep and deep sleep … What I eventually understood was that I was having not enough deep sleep”.

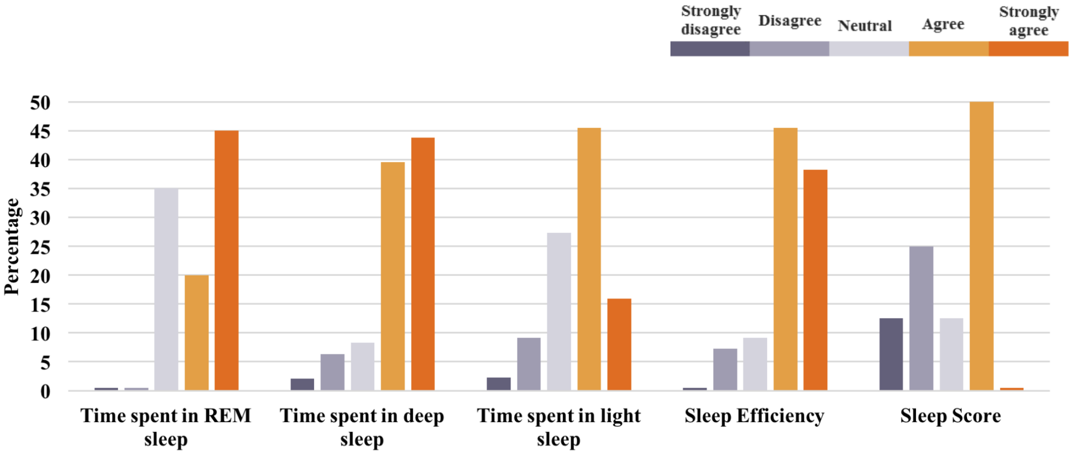


Figure 3.1: Percentage of survey respondents who believed their sleep quality was related to various sleep metrics

Second, there is limited research on how a person can take actionable steps towards affecting the number of hours spent in a particular sleep stage. “I’m not aware of anything necessarily that can increase REM sleep. Many medications, particularly psychiatric medications, can affect sleep architecture sum. What the effect that a medication would have on any given individual is there’s probably some variability to that” (E1). The feedback on sleep stages provided by commercial sensors promotes incorrect mental models on what these sensors can infer and how these stages actually impact sleep quality. “It’s … a lot of useless feedback… what I would like to see more clearly is really that the lay public understands there is no scientific basis for these numbers” (E5). Experts instead want feedback to focus on issues people actually have control over, such as sleep hygiene and modifiable behaviors. “That feedback might be if you’re not getting a lot of deep sleep, they might interpret that as a poor night of sleep and think that their getting bad sleep. Again, it doesn’t really lend itself to being actionable; so what are they supposed to do about that necessarily?" (E5). Experts expressed a desire to help users understand what these sensors can actually infer about sleep: "I think it [feedback from the device] needs to be scaled back into what we can expect them [users] to realistically understand and do something about" (E5).

### *3.6.4 Making unscientific correlations based on sleep stages*

For good sleep hygiene, avoiding large late night meals is recommended, but correlating with deep sleep stage might scientifically incorrect. Users may not have understood is that a big meal might cause restless sleep, and therefore cutting down on large meals might have allowed them to have less restlessness, leading to better sleep quality overall. In some cases, such inferences can sometimes lead to actions that can be potentially detrimental to health. For instance, users like R168 (Zeo) experimented with medication in an attempt to increase the duration of REM: "I can also see a day-to-day trend of how the iodine supplement I have just restarted is helping my sleep. It has increased my deep and REM and I feel better, even though my overall sleep time is not that much more."

## *3.7 Discussion and Design Recommendations*

Our findings show that sleep sensors increase awareness in prioritizing sleep and help users address modifiable behaviors and their sleep hygiene. On the other hand, current feedback focuses on sleep metrics people do not have control to directly change (e.g., time in sleep stages) and this distracts users from focusing on aspects they have control over that improve sleep health. We now provide design recommendations for on the feedback sleep sensing technology can provide to users. Our guidelines draw from our results, and connect to evidence-based strategies that focus on sleep hygiene, modifiable behaviors, and the SATED framework for good sleep quality.

### *3.7.1 Include Subjective Sleep Quality Assessment*

Sleep quality is inherently subjective. A poor night's sleep for one person can be satisfactory and rested sleep for another person. Furthermore, the effects of a poor night's sleep vary from person to person. Sleep quality self assessments is often used by clinicians to assess the severity of sleep-related issues [27]. We recommend that subjective self-assessments be incorporated as part of the analysis that sleep sensing technologies execute to calculate people's sleep quality for a given night. To assess subjective sleep quality, we recommend incorporating the five dimensions of the SATED framework, such as Satisfaction and Alertness with more objective measures such as Efficiency, Timing, and Duration. Incorporating users' subjective assessment should also be integrated into algorithms that personalize feedback or calculate a sleep score. Furthermore, self assessments should be used to learn and assess which types of modifiable behaviors worked best in helping a user improve their sleep over time.

### *3.7.2 Contextualize Sleep Quality with Journaling*

The current state of feedback does not support long-term perspectives on sleep trends. We recommend sleep technologies support long-term visualizations of bed time, wake time, and sleep duration. Long-term visualizations can provide a richer and more holistic view on variability compared to daily feedback focused on sleep stage. Viewing long-term trends will help users address aspects of sleep hygiene related to maintaining a consistent bedtime and wake time.

### *3.7.3 Focus on Actionable Feedback*

We find that feedback helps users connect their daytime behaviors, pre-bedtime behaviors, and environmental conditions of their bedrooms to their sleep quality, which in turn helps them act accordingly. This confirms previous research [31, 65]. To help people draw meaningful conclusions from sleep data, designs need to develop ways of presenting feedback to users beyond correlational graphs. Moreover, support for more systematic tests such as through self-experiments [42, 64], can make this process less frustrating than simple trial and error. Systems can allow people to test behaviors such as the timing of caffeine consumption or installing noise and light blocking curtains. Reviewing these experiences will help people identify the impact of that change on their sleep duration, timing, or satisfaction.

## ***3.8 Summary***

Sleep sensing technology provides people with rich information about their sleep. These technologies help people learn about their sleep habits and how to improve sleep health by providing feedback on their sleep. However, certain types of feedback lead users to develop broken mental models about what sleep sensors have the ability to sense and distract users from habits and behaviors that are actually affecting their sleep. Across different commercial sensors, the metrics used to give sleep quality feedback vary and sometimes conflict with clinical standards, potentially undermining people's ability to improve their sleep. The focus on sleep stages, which are difficult to infer from the set of sensors sleep technologies use, leads users to focus on aspects of their sleep difficult to control, such as REM sleep. This focus derails users from focusing on modifiable behaviors and sleep hygiene. Our findings provide a review of the state of current sleep sensing technology from the perspective of users and sleep experts. Tools should focus on actionable feedback that integrates modifiable behaviors. Sleep self-assessments can help personalize and contextualize sleep sensing feedback.

# Chapter 4

# CONTINUOUS HEALTH MONITORING: WIBREATHE

## *4.1 Introduction*

In this chapter I describe the design and evaluation of a 2.4GHz CW (continuous wave) wireless whole-home breathing monitoring system: WiBreathe. The WiBreathe algorithm, designed to provide accurate breathing rate estimation from a wireless base-band signal, shows promise for the reliable use of a non-contact system for physiological signal monitoring. The WiBreathe system was evaluated in controlled lab settings as well as in natural home settings with users performing activities. WiBreathe's algorithm was able to reliably detect breathing rate across different scenarios with an average error rate of 2.16 breaths per minute. It is notable that WiBreathe's error rate when participants were lying down decreased significantly to 0.96 breaths per minute. This chapter lays the foundation for a wireless signal based, unobtrusive sensing of sleep-related physiological bio-markers.

## *4.2 Motivation*

Continuous, non-invasive, and unobtrusive sensing of various health metrics has the potential to improve an individual's well being and quality of life. This information not only provides them with timely feedback on their overall physiological condition but also helps detect abnormalities in trends over prolonged periods of time. Of the many health metrics, detecting respiration rate in a home environment has significant impact in determining potential pulmonary exacerbation in advance. In general, respiration rate in particular is used as a physiological measure for tracking diseases in many areas, such as sleep, pulmonology, and cardiology, and can also provide useful insights about the psychological and psycho-physiological condition of an individual. In particular, continuous monitoring of respiration rate throughout the night is relevant for diagnosing as well as monitoring sleep apnea; an apnea event is when a patient is not breathing for 10 seconds or longer during sleep [110]. The breathing patterns associated with high stress levels can provide the user relevant information regarding their psychological well-being over long periods of time.

WiBreathe is a whole-home respiration rate sensing system that reliably estimates respiration rate in a user's changing environment. Our[1] system leverages wireless narrow-band signals to monitor the breathing of an individual anywhere in a home, even when the person is behind walls. Specifically, our algorithm clusters and chooses between multiple respiratory rate extraction algorithms, and adapts to a dynamically changing environment. Using a single transmitter-receiver pair (see Figure 4.1), we show the ability to detect breathing during various activities such as reading, typing at a desk, watching television, and lying down. Given the ubiquity of wireless signals (e.g., Wi-Fi), such an approach can enable true continuous breathing detection throughout the day at various locations in a home, obviating the cost and inconvenience of deploying multiple sensor systems around the home.

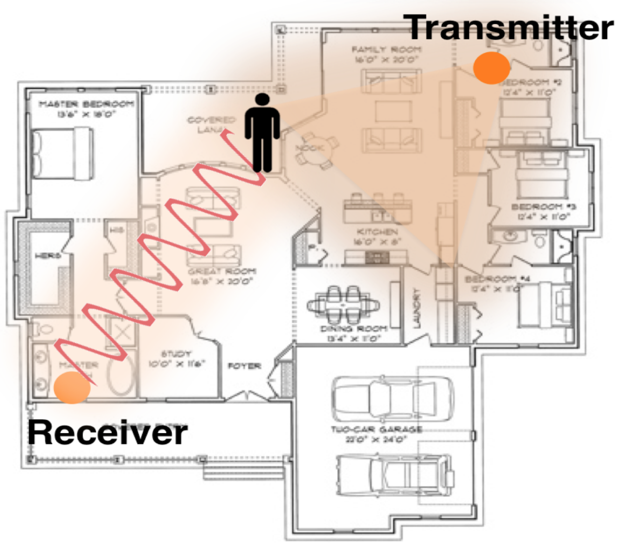


Figure 4.1: WiBreathe can detect a person's respiration rate from anywhere in a house without any instrumentation on the body. The system only requires a pair of transmitter and receivers that can be placed anywhere in the house.

---

[1] This study was conducted in collaboration with the following co-authors: Elliot Saba, Ke-Yu Chen, Mayank Goel, Sidhant Gupta, Shwetak N. Patel.

Although clinical devices have are not evaluated in natural settings or when users are performing activities, research has shown that they tend to have similar variability for respiration rate monitoring during sleep [33]. WiBreathe enables whole-home breathing detection that works through walls and over large distances using a single transmitter-receiver pair. WiBreathe's adaptive algorithms remains immune to changes in the environment and the user's breathing pattern.

## *4.3 System*

WiBreathe leverages 2.4 GHz wireless signals to detect the respiration rate of a person located anywhere in a home. To do so, the user places a transmitter-receiver pair at any location in the house without any calibration. Ideally, the transmitter and receiver are placed in the two corners of the room to provide full coverage (see Figure 4.1).

### *4.3.1 Theory of Operation*

The two phases of breathing: inspiration and expiration involve the intercostal muscles and the diaphragm. The expansion and contraction of these muscles cause a corresponding increase and decrease in the lung volume and by extension, displacement of the chest and abdomen. When wireless signals propagate in a medium, apart from the direct path between the transmitter and the receiver, the signal reflects off various objects in the environment. The human body, when present in the path of signal propagation, acts as a reflector. The minute chest and abdomen movements amplitude-modulate the wireless signal before it reaches the receiver. WiBreathe captures these signal variations and deduces the respiratory rate from the periodicity of the amplitude modulation. In particular, our system employs an envelope detection algorithm to extract the breathing frequency from the 2.4 GHz wireless signals.

### *4.3.2 Hardware*

We used two Ettus USRP (Universal Software Radio Peripheral) N210 to prototype the WiBreathe system as it allows access to a high performance FPGA that can be used to generate a continuous signal at 2.4 GHz. One USRP was configured as a transmitter and another as a receiver. In our work, we chose the XCVR2450 RF transceiver frontend that operates in the 2.4 GHz to 5.9 GHz range.

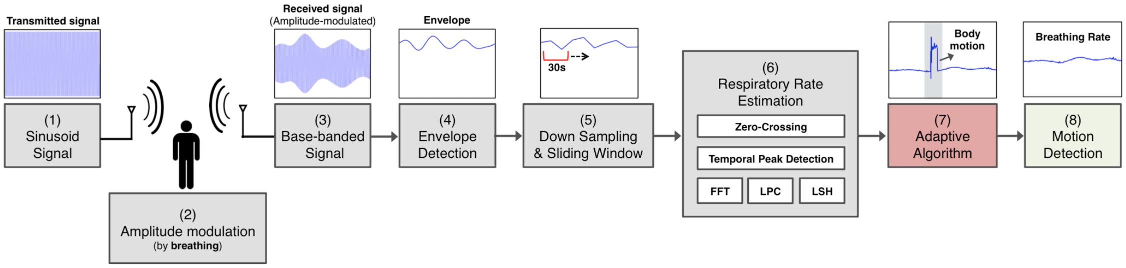


Figure 4.2: Signal processing for the breathing detection algorithm. After demodulating and detecting the modulation envelope, a 30-second sliding window is applied to compute the average breathing rate for each window using all five sub-algorithms (Step 6). These breathing rate estimates are input into an adaptive algorithm that selects the best estimate (Step 7). The harmonic fit from the LSH estimator is then used to classify windows containing body motion, which are ignored (Step 8).

The USRP generates the continuous signal with a center frequency of 2.4 GHz over a narrow bandwidth of 20 MHz similar to a single sub-channel of an Orthogonal Frequency Domain Multiplexing (OFDM) wireless signal. The transmit power is set to 15 dBm. In addition, LP0965 directional antennas with a gain of 6 dBi and antenna factor of 32 at 2.4 GHz are used to enhance the signal strength. In the following sections, we refer to the USRPs along with the directional antennas as access points (AP).

### *System and Implementation Details*

Figure 4.2 shows the data flow diagram of the WiBreathe system. We transmit a continuous sinusoid from the USRP transmitter at 2.4 GHz (Fig 4.2, step 1). Breathing alters the magnitude of reflected signal inducing an amplitude modulation on the transmitted wireless signal. (Fig 4.2, step 2). This amplitude-modulated signal can be mathematically modeled as:

$$u(t) = A_c m(t) cos(2\pi f_c t) \tag{4.1}$$

where Ac and fc are the amplitude and frequency of the 2.4 GHz carrier signal, respectively, and m(t) is the breathing signal acting as modulator upon the carrier. At the receiving end, the access point first demodulates the 2.4 GHz amplitude-modulated signal down to baseband and samples it at a rate of 32 KHz (Fig 4.2, step 3). The transmitted sinusoid is then narrow-band filtered and demodulated by taking the magnitude of a single DFT bin over overlapping windows across the entire signal, generating an envelope of the received signal (Fig 4.2, step 4). This rejects all spectral energy not centered around the transmitted sinusoid, keeping all data within 23 Hz of the carrier frequency. Since the respiration rate of a human usually does not exceed 1 Hz, we further down sample and filter the signal to restrict the frequency domain to contain only the relevant spectral energy (i.e., 0.2 - 1.0 Hz). We simultaneously reject high frequency noise as well as the strong DC component present in all recorded signals through this bandpass filtering. (Fig 4.2, step 5). At this point, the down sampled signal is similar to a carrier signal amplitude-modulated in the range of 0.2 to 0.5 Hz, (i.e., the human's normal breathing frequency) [100], and we can begin frequency estimation.

### *4.4 Algorithm*

To estimate the frequency of the breathing signal, we divide the signal into 30s sliding windows, with 97% overlap between them (Fig 4.2, step 5). Each window is then analyzed with multiple frequency estimation algorithms including (1) Zero-crossing detection, (2) Fourier transform maximum selection (FFT), (3) Linear Predictive Coding (LPC) and (4) Least-Squares Harmonic analysis (LSH) (Fig 4.2, step 6).

Zero-crossing detection attempts to directly estimate the frequency of a periodic signal by measuring the number of negative-to-positive transitions of a time waveform in a given time window. Temporal peak detection similarly measures the number of local maxima/minima of a time waveform in a given time window. Both of these methods work well in the absence of noise; however as noise increases, it is to be expected that their performance will suffer.

Fourier transform maximum selection (hereafter denoted as the FFT method) takes the DFT of a given time window, and simply selects the frequency with the largest component. This method was employed by [9] in a highly controlled setting and has shown promising results.

Linear Predictive Coding (LPC) analyzes a given time window by learning the linear relationship of samples in that window via a least-squares estimation method. This algorithm was first introduced by O'Shaughnessy,D. [51] and has found wide applications in fields such as speech processing. The linear relationship found through LPC represents a filter that estimates power spectral density of the signal, and therefore using the result of LPC analysis allows us to estimate the location of the dominant spectral shape in the signal being analyzed. Furthermore, this algorithm is not quantized in its accuracy as compared to methods such as zero-crossing, temporal peak detection, or the FFT method. In all these previous cases, the output value is quantized by either the temporal sampling rate, or the Fourier bin size; however LPC is capable of yielding frequency values with sub-bin resolution.

Least-Squares Harmonic analysis (LSH), proposed by Qin, L. [87], takes advantage of the fact that the breathing signal is not perfectly sinusoidal, but is still roughly periodic. Periodic non-sinusoidal signals create harmonics at integer multiples of the fundamental frequency, and LSH uses this fact to gain better results in situations with high noise. It does this by using the Goertzel algorithm to analyze spectral energy at specific, harmonically related frequencies, and chooses the fundamental frequency whose harmonics have the greatest total energy. The Goertzel algorithm uses a filter to determine the energy in a narrow band of frequency similarly to the DFT, but with an arbitrary frequency resolution, which we have chosen to be 0.001Hz between 0.1 and 0.7Hz. This addresses the quantization issues mentioned above, and exploits the fact that breathing signals have at least a second harmonic with significant energy above the noise floor. This allows LSH to gain sub-bin resolution similar to LPC, and aids the estimator in the presence of noise, because the extra energy in the second harmonic effectively increases the signal power.

These four algorithms each take advantage of different features of the breathing signal, performing very differently depending on factors such as the subject's breathing pattern, the ambient noise level and the orientation of the subject relative to the transmitter and receiver. In Figure 4.3 the waveform displays a large amount of high frequency energy due to a combination of user motion and noise. This high frequency energy causes the zero-crossing and LPC to severely overestimate the breathing rate resulting in an comparatively high error percentage of 51% and 67% respectively, whereas the Least-Squares Harmonic analysis and FFT estimator are able to accurately locate the base harmonic of the breathing signal, and give a closer approximation with 21% and 31% error rate respectively.

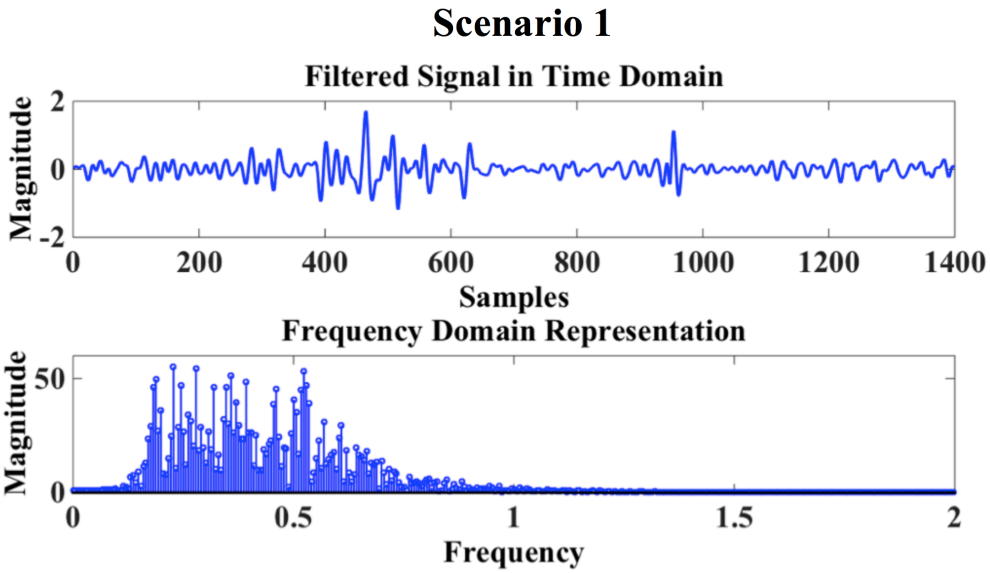


Figure 4.3: Breathing waveform with time and frequency domain representations. Significant high frequency energy due to activity and signal attenuation, reducing the performance of zero crossing and LPC.

Conversely, in Figure 4.4 is shown a situation in which the signal is particularly non-sinusoidal due to shallow breathing, and therefore lacks a strong single dominant frequency, preventing FFT and LSH estimators from accurately determining the breathing rate. The LPC and ZC methods on the other hand are better equipped to estimate the central frequency component in a large group of frequencies of similar magnitudes when spectral energy due to noise is not an issue, and therefore are able to perform with an error rate of 7% and 15%, as opposed to FFT and LSH estimators which result in 21% and 26% error respectively.

Our analysis shows that no single method is able to perform adequately across all situations since different factors induce varying frequencies close to the actual respiration rate. To address this, we have created an adaptive algorithm that dynamically combines and selects results from all four algorithms presented (Fig 4.2, step 7). Initially, the algorithm determines the mean of two closest frequency estimates between all the four algorithms. Once we have at least 5 past estimates, the adaptive algorithm then takes the results from all sub-algorithms for the current window and selects

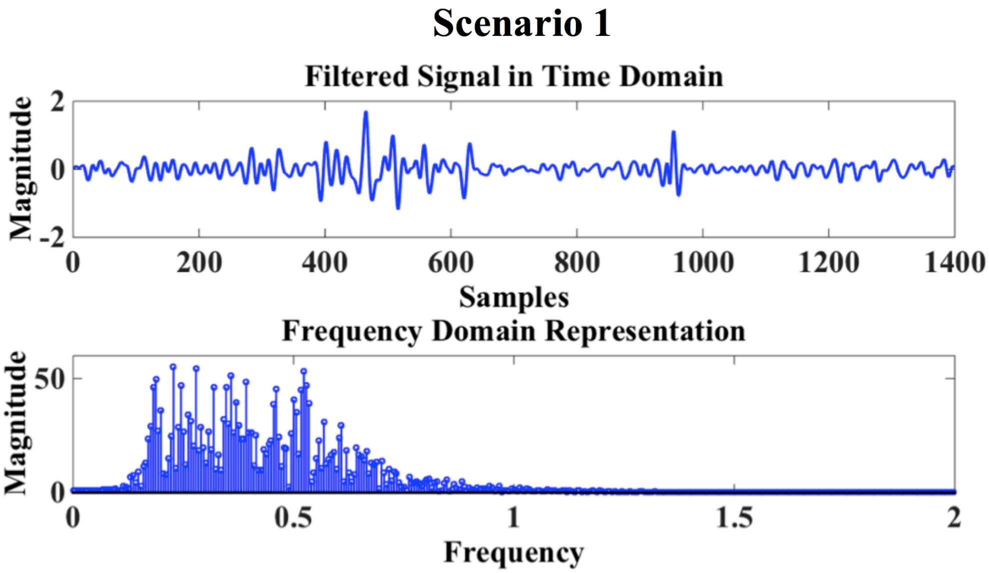


Figure 4.4: Breathing waveform with time and frequency domain representations. Non-sinusoidal breathing pattern lacking a dominant frequency causing increased error rate in FFT and LSH estimators.

the result closest to the median of the last five frequency estimates. All the four individual estimation techniques estimate a new value every 0.9 seconds, therefore using the past 5 estimates is a reasonable choice since the respiration rate does not change drastically over 4.5 seconds. Furthermore it assumes that at least two of the four sub-algorithms are able to estimate the respiration rate with moderate accuracy in any scenario. This way, our system is able to select the best sub-algorithm such that the frequency estimate does not vary dramatically over short time periods, a phenomenon which plagues all of these five frequency estimation methods and results in large spikes in error rate. Combining the results across multiple algorithms to adaptively choose the best estimate, allows this algorithm to outperform any single algorithm by avoiding these large deviations from the true respiration rate.

### *4.5 Evaluation and Results*

We evaluated the performance of WiBreathe in four different conditions and compare the performance against the individual techniques mentioned in the System and Implementation Details section. Apart from the four techniques used in WiBreathe, we also evaluate the results from another common technique Peak detection, used mainly for breathing detection in wearable or contact breathing detection. For all the experiments, participants wore a Vernier respiration-monitoring band around their abdomen to capture the ground truth respiration rate. The strap is similar to a brachial blood pressure cuff and consists of a bladder that is filled with air. Since the outward and inward abdomen movements cause changes in the air pressure, a differential air pressure sensor is used to convert the varying air pressure to a voltage ranging from 0 to 5V. The analog voltage was then recorded using a data acquisition unit. We conducted experiments in two settings. The first set of experiments was done in a controlled lab environment and the second set of experiments was performed in participant's homes in a natural setting.

#### *4.5.1 Controlled Settings*

To explicitly evaluate the trade off between increased SNR and coverage of the directional antennas we conducted a controlled study with 3 participants in an office space. Participants were asked to sit at various positions with respect to the transmitter receiver pair and their breathing rate was recorded for a period of 3 minutes in each position.

**Effect of Distance**: Participants were asked to sit at 5 different positions with a distance of 0.9m, 1.5m, 2.1m, 2.7m, 3.3m, 4.3 m. Fig. 4.5(left) shows the error rate for all the individual algorithms as well as WiBreathe's combination algorithm when the participant sat at various distances from the transmitter-receiver pair.

We can notice a general trend in error rate for all the algorithms: the error increases considerably when the participant sits closer to the transmitter-receiver pair (i.e., less than 1.5 m). The reduced performance when the user is too close to the antennae results from the directionality of the antennas. As the directional antennas used in the experiment have a radiation pattern similar to that of a cone radiating outwards, the intensity of the signal is higher along the main-lobe while slightly attenuated along the side lobes. Therefore, the positions located in the range of the main-lobe (i.e. 1.75m

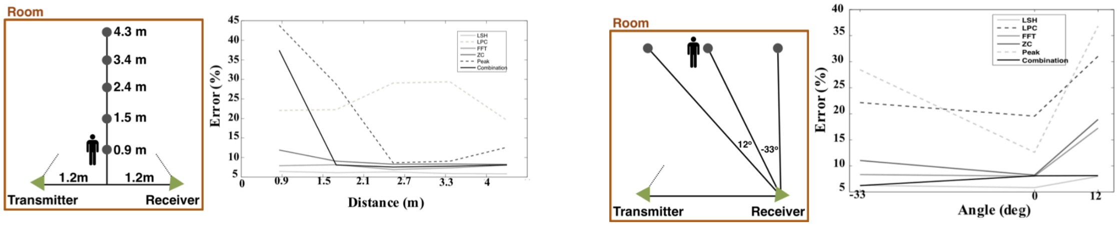


Figure 4.5: Experimental test-bed for Study 1: Line-OfSight. (Left) Data was collected at 5 locations (marked as grey points) of different distances. Error percentage (%) for Study 1: Line-Of-Sight (various distances) of four estimation techniques and combined WiBreathe algorithm. Experimental test-bed for Study 2: Line-OfSight. (Right) Data was collected at 3 spots with various orientations (incident angles of 12o, 0o and -33o) relative the receiver. Error percentage (%) for Study 2: Line-Of-Sight (various orientation), of four estimation techniques and Combined (WiBreathe) algorithm.

4.3m) yield a low error rate while in the positions that fall out of range of the main lobe have higher error rates. In case of the Combination algorithm, this variability in error percentage due to distance is reduced and the performance of the system stabilizes. The fact that the Combination algorithm performs better even when close to the antennae suggests that WiBreathe is able to counter the variability of signal strengths and a user's positions. It can be used reliably in most locations and is less prone to dead-zones.

**Effect of Orientation**: Orientation: To study the effect of orientation of the individual with respect to the antennas on the respiration rate in a controlled environment, we recorded data at three different positions as shown in Fig. 4.5(right). We can observe slightly increased error percentage in frequency estimation for individual algorithms in case of off-center orientations. When the angle of incidence is equal to the angle of refection (e.g.. 33o), there is total internal reflection like conditions for the signal, making the error rate comparatively lower. As the user moves away from the point of optimal reflection, the intensity of the reflected signal decreases resulting in decreased signal to noise ratio. This leads to higher error rates. However, by adaptively choosing the best estimates from the five algorithms, we are able to resolve this problem and make the system more reliable with change in position. Our approach has an average error of 9.8% when the user is at the center and 7.1% when the user is off-center, as compared to 7.5% and 15.25%, respectively, without using

WiBreathe's combination of algorithms.

### *4.5.2 Natural Settings*

In order to evaluate how well WiBreathe works in a natural setting, experiments were conducted across 4 homes varying in square footage from 600 sq. ft. to 2000 sq. ft. with 6 unique participants, ranging from 23 to 40 years old (3 female). Each home had at least 6 LOS and 2 NLOS scenarios. Prior scientific literature and our controlled experiments have both suggested that a number of factors could affect the performance of a wireless sensing system in an environment. Of those, the most relevant and crucial ones are: 1) Environmental factors such as distance and orientation of the user from the antennas and 2) Subject variations such as movement, posture and activity of the user. To test the performance of WiBreathe across these factors, we collected data for each participant at multiple locations while they performed different activities. The users were asked to sit at different positions in a room where the antennas were placed in two corners as shown in Figure 8. The extracted breathing signal from WiBreathe and the ground truth was captured for a period of 3 minutes in each position.

**Distance and Orientation**: The directional antennas used in our experiments were placed in locations that maximize the sensing coverage of users' daily activities such as reading, watching tv, or typing. At each positions, we examined the effect of distance and orientation between the user and the antennas.The six positions varied in distance from 1.5 m to 6 m across 4 homes. Figure 4.6(left) shows the error rate for all of the individual algorithms as well as WiBreathe's combination algorithm when the participant sat at various positions with respect to the antennas. Our approach presents an average error of 1.92 breaths per minute (bpm) across all 6 positions.We can also observe that the variablity in error rate is higher for all other individual algorithms compared to WiBreathe across the six positions.The observation that the WiBreathe algorithm performs better across all positions suggests that WiBreathe is able to counter the variability of signal strengths and user orientations by adaptively choosing the best estimate from the individual algorithms. It can be used reliably in most locations and is less prone to dead-zones instead of using any of the other single techniques.

**Activity**: We also evaluated the performance of WiBreathe when we factor in motion and prob-

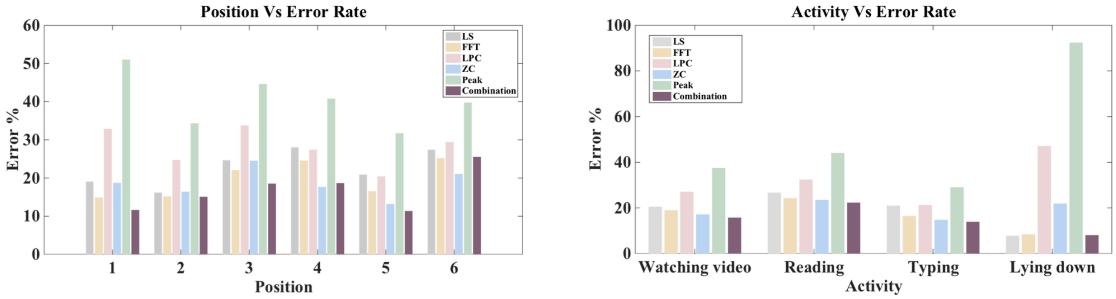


Figure 4.6: Error percentage (%) for Study 1: Line-Of-Sight (various distances) of five estimation techniques and our approach: LSH, FFT, Zero Crossing, Peak Detection and Combination.

able changes in posture when users perform everyday activities. The users were engaged in four different activities, in decreasing order of intensity of motion: typing, reading, watching a video, and lying down. In order to simulate sporadic body movements while performing these activities, like stretching, adjusting their posture, etc., we instructed the users to move at timed intervals. This served as ground truth since we could determine when the algorithm ignored segments from motion and not due to addition of noise by other environmental factors. While it is expected that activites such as typing would have a higher error rate due to increased motion, Figure 4.6(right) shows that across the three activities where the participant is sitting, reading has a slightly higher error rate. Amongst the various factors that contribute to error, we have observed that the participant's posture (sitting straight versus slouching), induce much more noise in the signal compared to finger and hand movements. This effect is further shown by the fact that the error is least, 0.96 bpm, when the participant is lying down.

### *4.5.3 Breathing Pattern Detection*

Our detailed discussions with medical professionals informed us that for most medical applications, the detecting changes in a user's breathing pattern and respiration rate are far more important than determining respiration rate when users are engaged in motion. For example significant variance in respiration rate of an individual or repetitive shallow breaths can point towards respiratory distress or dyspnea. Detection of time periods where there is no breathing signal or discontinuities in the breathing signal point to apnea and arousal events during a sleep cycle which helps in diagnosis of sleep apnea. This means a system that adapts to sudden variations in breathing rates can be equally or even more useful than the earlier systems. Therefore we investigated how WiBreathe performs if a user's breathing pattern changes considerably over a short duration of time. Figure4.7 (top) shows the received signal versus the ground truth in a situation where the user is initially breathing slowly, then suddenly increases their respiration rate. This period of rapid breathing is then followed by a period of no breathing, then a short time of deep slow breaths followed by a few fast and slow breaths. The performance of various algorithms and WiBreathe's combined algorithm can be seen in Figure 4.7 (bottom). Although WiBreathe rejects sudden changes in frequency in a short window, when multiple algorithms show consensus it is able to get a closer estimate across the entire time frame resulting in reduced average error rate.

More experimental set-ups and results can be found in the paper [106]

## *4.6 Discussion*

### *4.6.1 Dynamic Algorithm Selection For Higher Accuracy in Natural Settings*

From our results, we observe that some algorithms have high resilience to distance and orientation such as zero crossing and FFT with an average error percentage of 18.5% and 19.6% respectively while some adapt better in the presence of noise due to activity such as FFT and LS estimator with error percentages of 16.8% and 18.8% respectively. Overall there is no single algorithm that works the best in all situations. Given these results, we designed WiBreathe to adapt to the natural variations in breathing frequency as well as changes in the environment. Our approach combines the estimates from an ensemble of four different techniques and adapts dynamically over a different scenarios to enable continuous respiration rate monitoring over long periods of time.

### *4.6.2 Implementation in Commodity Wi-Fi Routers*

Given that our preliminary experimentation suggests that WiBreathe's respiration rate estimation technique is reliable under various conditions, we believe the next step is to apply similar signal processing and frequency estimation techniques presented here to actual Wi-Fi signals. This would make the system truly ubiquitous and allow a Wi-Fi router to receive RF signals from any single

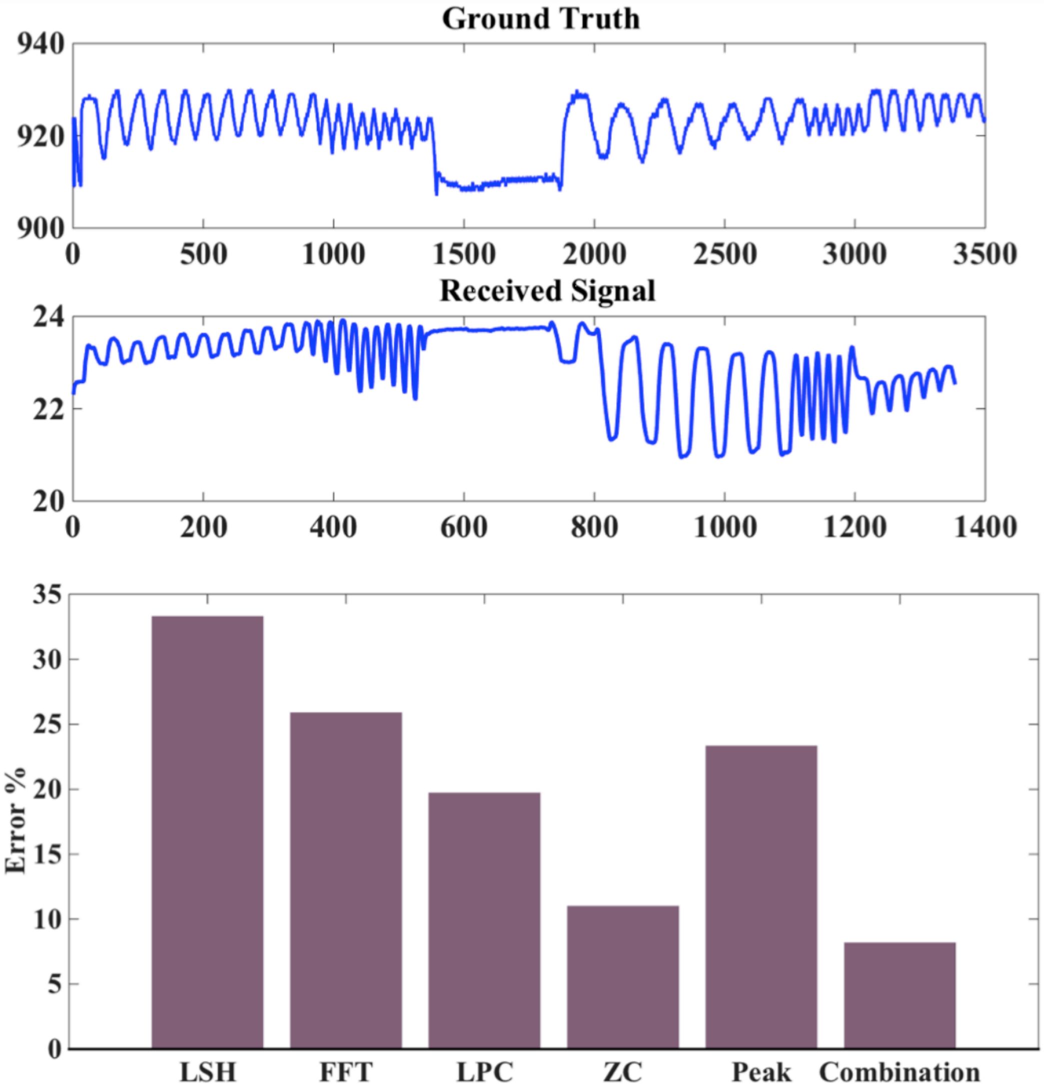


Figure 4.7: (Top) Breathing pattern detection in reflected raw received signal. (Bottom) Performance of various algorithms for drastic changes in breathing rate.

Wi-Fi transmitting device (Access point, cellphone, laptop, smart TV, etc.) for respiratory rate estimation. The directional antennas used for WiBreathe provide significant gain and directionality and aid in breathing pattern detection. However, omnidirectional antennas used in commercial routers may not be able to provide such high gain. The adaptation of the current algorithm to an actual Wi-Fi router can be explored further.

### *4.6.3 Vital-signal Detection of Mutiple People*

In order to verify the feasibility of using WiBreathe to detect the respiration rates of multiple individuals present in the same room, we conducted a preliminary study with two users. The participants were asked to breathe to a metronome at set frequencies. Figure 13 shows the FFT plots with the estimated frequencies of the two individuals. The relative position of the users and the breathing patterns causes one of the frequencies to be dominant. By implementing a simple notch filter we can filter the dominant frequency and the second person's frequency can be then estimated. We see that for multiple users the sub-algorithms can reliably detect respiration rate with an error of 1.54 bpm. While frequencies of two people breathing in the same environment can currently be estimated given certain preconditions, a more useful application of the system would be the ability to track the respiration rate of a single person in the presence of multiple people. This would be helpful in the case of tracking the respiration rate of one person when two people are sleeping side by side on a bed. This can be implemented by beamforming, specifically by leveraging MIMO systems available in modern routers to estimate the Angle of Arrival of two signals using Multiple Signal Classification.

## ***4.7 Summary***

In this chapter, I have presented a non-invasive breathing monitor system that requires no instrumentation on the human body. Using a single pair of transmitter and receiver, our system is able to monitor the respiratory rate of an individual located anywhere in the house. Our algorithm takes the results adaptively from five different sub-algorithms, making the system robust in a dynamically changing environment. We evaluated the system in both a lab-controlled and real home environment. The results show an average error rate of 2.16 bpm in a natural setting across our 6 participants, which is comparable to a clinical breathing monitor. As health sensing becomes more commonplace, WiBreathe enables whole-home, continuously respiratory rate monitoring, obviating the need for any wearables on body.

Through WiBreathe system we have demonstrated the feasibility of using wireless signals for continuous non-contact physiological signal monitoring in a home environment. In the next chapter we can see how the wireless signals can be used for sleep sensing.

# Chapter 5

# CONTINUOUS SLEEP MONITORING: DOPPLESLEEP

## 5.1 Introduction

In this chapter I describe the implementation and evaluation of a 24GHz FMCW (Frequency Modulated Continuous Wave) RADAR system: DoppleSleep; for sleep estimation. DoppleSleep shows promise in being able to unobtrusively sense three sleep related biomarkers: breathing, heart rate and body movement. By extracting 42 frame level features from these three signals, we[1] then classify sleep versus wake state and derive sleep related objective metrics. DoppleSleep paves the path for an unobtrusive sleep sensing solution that can provide objective sleep metrics to aid in long term sleep tracking.

## 5.2 Motivation

An overview of the state of art in sleep sensing can inform us of the existing gaps in the practical usage of sleep sensing systems. At one end of the spectrum is polysomnography (PSG), which is regarded as the medical gold standard for assessing sleep quality and for diagnosing sleep-related disorders such as sleep apnea [62]. By instrumenting patients with at least 7 different sensors and electrodes that track various sleep-related physiological parameters throughout the night, PSG provides fine-grained sleep quality assessment. PSG is considered a highly obtrusive sleep sensing system due to its expense, impracticality for home-based use, and comfort-level for the patient, and thus its application is limited to diagnosing sleep-related disorders in clinical settings for a short duration. At the other end of the spectrum are commercial devices that enable long-term sleep monitoring. These primarily use Actigraphy, a widely adopted method that infers sleep duration and quality by measuring body motion during sleep [43]. Numerous consumer electronic fitness trackers [3, 8] leverage inbuilt accelerometers to provide actigraphy based sleep monitoring. However, many

---

[1] This study was conducted in collaboration with the following co-authors: Tauhidur Rahman, Alexander T. Adams, Mi Zhang, Shwetak N. Patel, Julie A. Kientz, Tanzeem Choudhury.

users are resistant to the idea of having to wear or place sensors close to the body during sleep [29]. To facilitate non-contact sensing, smartphone apps have been developed that sense sleep-related environmental factors in addition to body movement to infer sleep quality. A low-cost, long-term, contactless sleep sensing system that monitors sleep related physiological variables similar to PSG but with the unobtrusiveness and potential for long-term monitoring of actigraphy could bridge the gap between effectiveness and daily usage of sleep sensing.

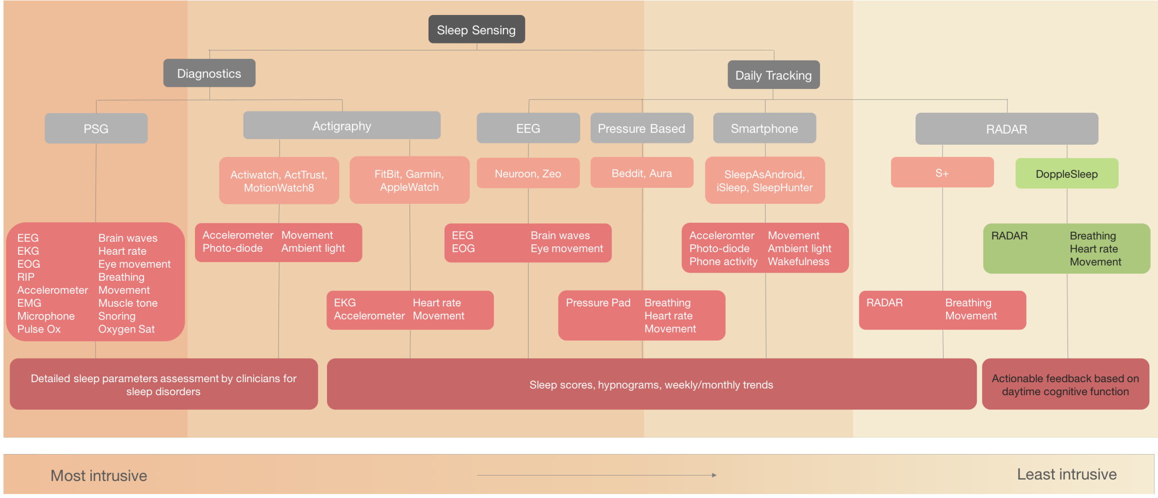


Figure 5.1: Comparison of DoppleSleep with a few state of the art sleep sensing technologies in terms of their typical usage, intrusiveness, affordances, sensing modalities, captured sleep biomarkers and predicted sleep variables.

DoppleSleep, is a contactless sleep sensing system (Figure 5.1) shows the that monitors three significant sleep- related biomarkers: breathing rate, heart rate, and body motion using a single, low-cost, off-the-shelf, K-band 24GHz radar module. Using a radar transceiver, the system tracks phase changes in the reflected electromagnetic waves and tracks the sleeper's body and limb movements. In the absence of large body movements, it also estimates the sleeper's breathing and heart rate using the periodic phase changes of the reflected wave from the expansion and contraction of heart and chest wall. DoppleSleep's breathing and heart rate estimation algorithm is relatively robust at various orientations and distances up to 2m between the user's body and sensor. It can estimate heart and breathing rate with an overall mean absolute error of 1.98 and 3.29 cycles per minute respectively. DoppleSleep uses these three sleep biomarkers to classify an epoch (the time unit for sleep classification) as a sleep or wake event with a recall of 89.6%. The sleep events are further classified as REM or NREM sleep stages with a recall of 80.2%. Lastly DoppleSleep objectively quantifies sleep quality using validated measures like sleep onset latency, number of awakenings, total sleep time, and sleep efficiency, similar to those produced in a PSG report [16]. DoppleSleep is comprised of an embedded system unit and a smartphone application unit. The embedded system unit locally samples and amplifies the raw radar base-band signal and transmits the signal to the smartphone via Bluetooth for further processing. The smartphone application unit is then used for heart rate, breathing rate, movement estimation, and sleep modeling. Although contactless detection of vital signals has been explored and refined over the last several decades, this is the first attempt that uses vital signals for sleep stage mining. DoppleSleep is a novel and complementary approach to sleep measurement that does not require contact with the user or the user's bed. We evaluated the contactless sensing of physical movements, heart rate, and breathing rate using short-range Doppler radar in both laboratory and real world settings. WE also developed and preliminary validation of Sleep vs. Wake and REM vs. Non-REM classification using objective sleep biomarkers and predicted sleep variables from sleep quality measurements on about 110 hours of sleep data collected from 16 sleep sessions with 8 participants.

### *5.3 Algorithm*

DoppleSleep explores the feasibility of contactless vital signal sensing using RADAR in the sleep-sensing domain. We selected a K band (24 GHz) direct conversion quadrature RADAR module, as it was most suitable for our application. The fundamental principle behind detecting vital signals using continuous wave (CW) Doppler radar is demonstrated in Figure 5.2.

The module transmits a single tone T (t) on a carrier frequency of f , wavelength of $\lambda$ = c/f , combined with phase noise $\phi$(t) from the oscillator, given by:

$$T(t) = cos(2\pi ft + \phi(t)) \tag{5.1}$$

Assume that T(t) traverses a distance of d0 and hits a human's body generating periodic chest movements due to respiration and heart beating. If the displacement of chest due to respiration is

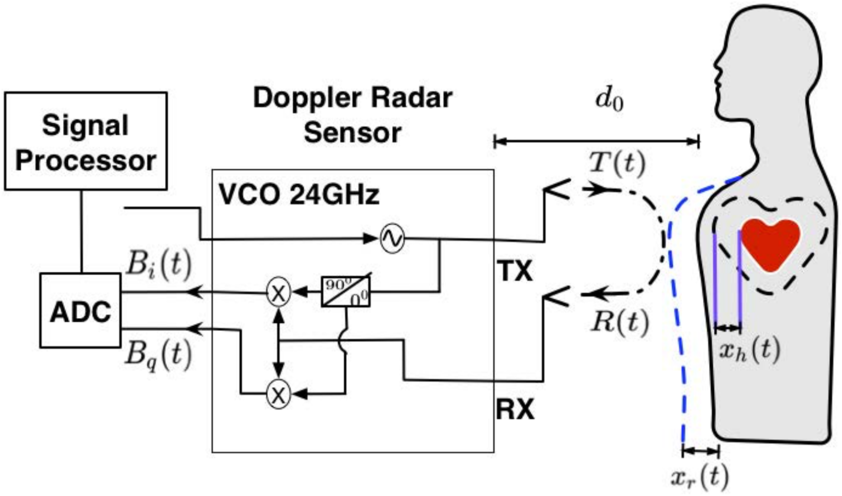


Figure 5.2: Detection theory of heartbeat and breathing using continuous wave (CW) Doppler radar.

$x_r(t)$ and the displacement of heart due to heart beat is $x_h(t)$, the overall movement can be expressed as $x(t) = x_r(t) + x_h(t)$. As a result, the reflected signal R(t) received by the radar is given by:

$$R(t) = A_r cos(2\pi ft - \frac{4\pi d_0}{\lambda} - \frac{4\pi x(t)}{\lambda} + \phi(t - \frac{2d_0}{c})) \tag{5.2}$$

R(t) is a time delayed and amplitude reduced version (reduced to $A_r$) of the transmitted signal T(t). Most importantly, the information of x(t) is phase modulated in R(t) in addition to the distance between the human body and the radar, $d_0$ and a time delayed version of the phase noise $\phi(t\frac{2d_0}{c})$. After R(t) goes through a Low Noise Amplifier (LNA), it is converted to baseband by a mixer that multiplies the received signal with a copy of the transmitted signal. The output of the mixer gives the difference or intermediate frequencies (IF). The receiver thus gets rid off any information related to carrier frequency $(2\pi ft)$ and preserves the change in phase of the signal corresponding to x(t) which we want to capture. In this study we use a quadrature receiver, which compensates for null detection points a problem faced by single channel receivers [46]. In a quadrature receiver, R(t) is split into two components and multiplied by two copies of transmitted signal that are 90 deg out of phase with each other. The output is thus a pair of orthonormal baseband signals, $B_I(t)$ and $B_Q(t)$, expressed by equation 5.3, 5.4.

$$B_i(t) = cos(\theta + \frac{\pi}{4} + \frac{4\pi x(t)}{\lambda} + \Delta\phi(t)) \tag{5.3}$$

$$B_i(t) = cos(\theta - \frac{\pi}{4} + \frac{4\pi x(t)}{\lambda} + \Delta\phi(t)) \tag{5.4}$$

Here, $\theta = \frac{4\pi d_0}{\lambda} + \theta_0$, contains the target distance information $d_0$ and $\Delta\phi(t)$ is the residual oscillator phase noise. The portion of interest is therefore the phase modulation due to physical and physiological movements x(t) given by $4\pi x(t)$. Since $B_I(t)$ and $B_Q(t)$ have a $90_{deg}$ phase difference,$lambda$ the quadrature receiver ensures that at least one of the baseband channels is not at a null detection point [46]. For example, if the distance d0 that makes up $\theta$ is such that $\theta$ is $\frac{\pi}{4}$, then $B_i(t)$ and $B_q(t)$ can be approximated as

$$B_i(t) = \frac{4\pi x(t)}{\lambda} + \phi(t) \tag{5.5}$$

$$B_q(t) = 1 - [\frac{4\pi x(t)}{\lambda} + \phi(t)]^2 \tag{5.6}$$

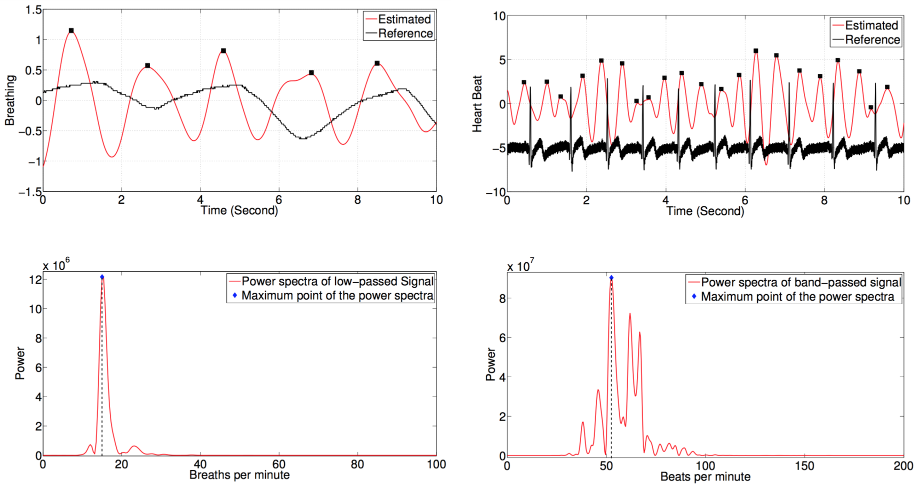
Here $B_i(t)$ is at an optimal point with full sensitivity, while $B_q(t)$ is at a null point with least sensitivity. Thus phase information can be recovered from one channel even if the other is at a null point. The next step is to process the two channels to get a output signal that is compensated for null-point. There are various null-point compensation techniques including frequency tuning technique [120], complex signal demodulation [81] and arctangent demodulation [102]. In this study we used a simpler technique of selecting one optimal channel that is farthest from the null point using interquartile range. Higher interquartile range of a channel will indicate that it is further away from the null-point, thus the optimal channel.

## *5.4 Sensing Physical Movements*

During sleep, out body manifests voluntary body movements such as tossing and turning, changing posture and involuntary limb movements such as myoclonic twitches [113]. The frequency and extent of the physical movements and vital signal variations during sleep can be used as indicators of sleep quality. Now we explain the algorithm used to segment physical movements during sleep. The challenge of using Doppler radar to track physical movements is that we must be able to isolate noise due to vibrations from appliances such as fan, air-conditioning unit or a speaker within the radar's range from human body movements. In order to address this challenge, we recorded physical activity data using the radar module and an accelerometer in our lab with 4 subjects. We then simulated three scenarios: (a) no physical movement, (b) common sleep related body movements (e.g. leg movements, tossing and turning, sitting up, head movements) and, (c) environmental noise induced by appliances. This was used as training data for our motion classifier. We then applied a low pass filter with cut off frequency at 3Hz, on the baseband signal, to remove high frequency periodic noise caused by environmental factors. As a result the frame-level RMS energy of the filtered baseband signal mostly corresponds to the presence or absence of body movements. However, there may be some frames where relatively high RMS energy may be caused due to aperiodic changes in the machine (e.g when the machine switches). In order to isolate these frames, the zero-crossing rate and the RMS energy of the filtered baseband signal are used as features for every 30 second frame. A leave-one-subject-out cross-validation experiment with a very simple threshold-based classifier indicates that these two features extracted in a frame level can easily discriminate among the three categories with an average recall of 94.5%. Using our algorithm to detect human movement, we found that the frame-level RMS energy of the filtered baseband signal is correlated (mutual information 0.86) with the frame-level RMS energy of the norm of 3D acceleration values from the accelerometer worn by our participants.

### 5.5 Sensing Breathing and Heart Rate

Once the signal frames containing movement data have been classified and isolated, we then proceed to estimate breathing and heart rate on frames that contain no body movement. As explained previously, the baseband signal Bi(t) is a linear combination of movement caused by breathing and heart beat signal $(x(t) = x_h(t) + x_r(t))$. As the breathing process generates relatively lower frequency signals than the heart beating process, we used two bandpass filters to isolate $x_h$ and $x_r(t)$ from $B_i(t)$. The bandpass filter for estimating breathing rate captures the lower frequencies that gets created in the baseband signal due to chest expansion and contraction. Specifically, we used a minimum order Butterworth filter with stop-band frequencies at 0.1 Hz, 0.8 Hz and passband frequencies at 0.3 Hz, 0.7 Hz to estimate any breathing rate ranging between 9 and 20 Breath per Minute. Similarly another minimum order band-pass Butterworth filter was designed with stop-band frequencies at 1 Hz, 3 Hz and passband frequencies at 1.5 Hz, 2.5 Hz to estimate any heart rate between about 45 and 80 Beat per Minute. For both filters, passband ripple and stop-band attenuation was chosen to be 1 dB and 60 dB respectively. These Butterworth filters have proven to be useful for vital sign estimation in prior literature [92]. These two filters are applied on the baseband signal $B_i(t)$ to get the estimated breathing and heartbeat waveform. Figure 5.3(a) shows the estimated and reference (or ground truth) breathing waveform from a respiratory inductance plethysmography (RIP) band. Similarly, figure 5.3(b) shows the estimated heartbeat waveform overlaid with the reference heartbeat signal from an electrocardiogram (ECG).



(a) (b)

Figure 5.3: (a) (Top) Estimated and the reference breathing waveform, (Bottom) Power spectral density of the estimated breathing waveform. (b) (Top) Estimated and the reference heart beat waveform, (Bottom) Power spectral density of the estimated heart beat waveform.

In figure 5.3(a), the peaks and troughs of the reference signal correspond to inhalation and exhalation. Notice that the estimated breathing waveform has peaks in both peaks and troughs of the reference signal. Hence, for every cycle of the reference breathing signal, we get about two cycles of our estimated breathing signal, one for inhalation and the other for exhalation. Thus the estimated breathing rate is half the frequency of the filtered signal. By applying a Fourier transform on the filtered baseband signal, we can then estimate the power spectral density of the signal as illustrated in figure 5.3(a)(bottom). The peak of the power spectral density corresponds to the dominant frequency, which is in this case the breathing rate. Similarly in figure 5.3(b), for every cycle of the reference heart beat signal, we get two cycles of our estimated heart beat signal. Thus the estimated heart rate is half the frequency of the filtered signal. Figure 5.3(b)(bottom) shows the power spectral density of the estimated heart beat, where the peak corresponds to the dominant frequency, which is the heart rate. In order to smooth our heart and breathing rate estimation, we applied a moving average filter with a length of 16. We also found that the window size of 30 seconds and shift of 5 seconds minimizes the heart rate and breathing rate estimation error.

For a sleep sensing system to work in real-world settings the vital signal tracking needs to be robust to relative orientation and distance between the radar sensor and the users' bodies during sleep. In order to evaluate the performance of DoppleSleep at various orientations and distances, we conducted two studies in a controlled environment. Four participants were recruited and their breathing and heart rate was recorded using the radar modules placed at different distances and orientation. The reference heart and breathing rates was recorded from a biometric shirt Hexoskin [5], embedded with an electrocardiogram (ECG) to provide heart rate value and respiratory inductance plethysmography (RIP) to provide breathing rate value. Comparing the estimated heart and breathing rates to the reference values we computed the mean absolute error.

### *5.5.1 Distance Test*

The distance d0 between the subject and the radar is directly correlated to the power of the reflected R(t) and the baseband signals $B_I$ , $B_Q$. Specifically, as the distance $d_0$ increases, the reflected signal gets weaker ($A_r$ß0 in equation 5.2) due to signal loss when propagating through the longer distance. The aim of this test was to find out the effect of distance on DoppleSleep's heart and breathing rate detection algorithm. We recruited 4 subjects and asked them to lie down in a supine position in our laboratory. We then varied the distance between the Radar and the subject from 0.5m to 2m with

0.5m increments. The overall error rate stays within 2 CPM (cycle per minute) for distances up to 2m.

### *5.5.2 Orientation Test*

The orientation test was conducted to explore the effect of relative orientation of the radar sensor with respect to the user's body, on heart and breathing rate estimation. Different orientations allow the sensor to capture different profiles of the human body as the heart and chest wall compress and expand. Thus different orientations may result in different error rates for heart and breathing rate estimation. We estimated heart and breathing rates using the radar at five different orientations with respect to the subject's body: (A) facing the bottom of the feet, (B) facing the chest, (C) facing the side of the torso, (D) facing the top of the head and (E) facing the back. In all the five orientations the radar was 1 meter apart from the subject's body. The lowest error rate for heart rate and breathing rate estimation is achieved when the radar faces the back of a subject's body. In other words, if user places the sensor underneath the bed, the performance of heat and breathing rate estimation might be maximized. Our result is in accordance with the findings in [83] where the authors explained that the accuracy of heart rate and breathing rate estimation from the back is maximized due to the minimal harmonic interference. Also we can observe that orientation A yields relatively smaller error rate. This could be attributed to the fact that the radar captures a larger profile of the abdomen than the chest, which exhibits more motion due to breathing and heartbeat in a supine position. Although different orientations yield different error rates, the overall error rate is within 3 CPM (cycles per minute) no matter where the sensor is placed.

## ***5.6 Evaluating DoppleSleep in the Wild***

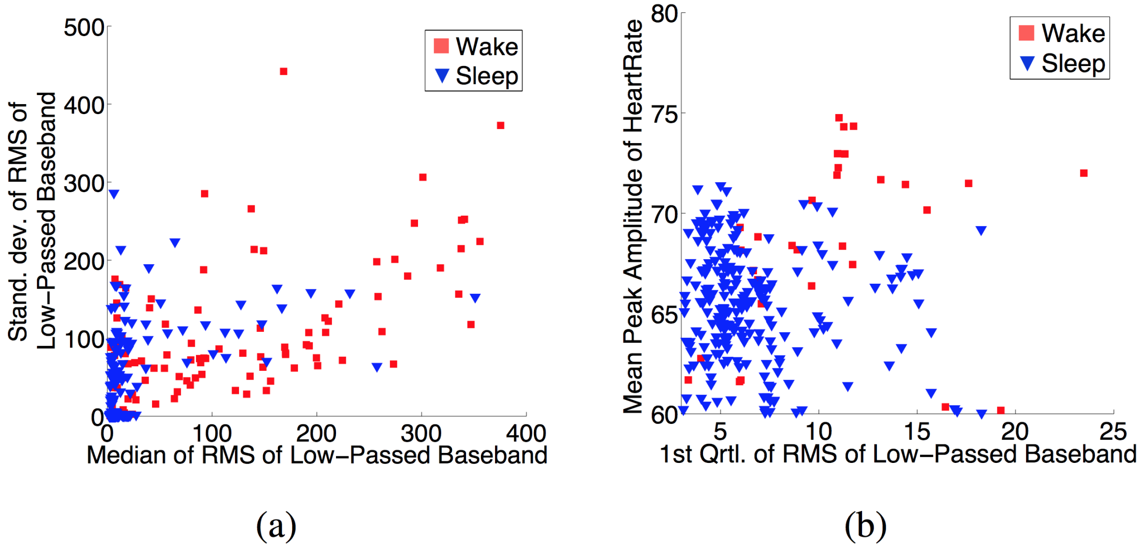
We recruited 8 healthy participants with no prediagnosed sleep disorders and collected sleep data for 2 sleep session each during their normal sleeping schedules at their homes. In total, we collected around 110 hours of sleep data. A biometric shirt (Hexoskin [5] ) was provided to capture ground truth heart rate using embedded EKG electrodes, breathing rate using respiratory inductance plethysmography, and physical movement using accelerometer. Two commercially available sleep-sensing systems were provided to track sleep stages and serve as reference for sleep quality parameters. Zeo

[15] is one such system that uses a headband embedded with EEG electrodes to track brain activity. Respironics [9] is another actigraphy-based system that predicts sleep stages based on accelerometer data. Both systems have been used in the research community as reasonably accurate references for wake/sleep and objective sleep quality parameters [54]. Participants were instructed to place the Doppler radar sensor at least 0.5 meters away from the body. The radar module's sensitivity in horizontal and vertical direction is respectively 80 degree and 34 degree. As long as the body is within the radar's angular coverage, the radar can effectively capture any movement from the body.

### *5.6.1 Sleep Modeling*

The sleep inferences (sleep vs. wake state and REM vs. Non-REM stage) from the EEG headband (Zeo) are used as ground truth for modeling sleep. Our sleep modeling starts with a Sleep vs. Wake classifier, which is a primary requirement for any daily sleep tracking purposes. Once a particular epoch of data is identified as sleep, we then classify the epoch into two stages: REM and NREM (or Non-REM). Lastly we objectively estimate sleep quality using clinically validated sleep quality parameters.

In order to train the Sleep vs. Wake classifier, we extract high-level features of the sleeper's physical activity, heart rate and breathing rate for every 5 minutes to predict sleep or wake states. We then apply different statistical functions to summarize various aspects of the activity, heart and breathing rate values in a frame. These statistical functions include extremes (min, max), averages (mean, RMS, median, quartiles), dispersion (standard deviation, interquartile range), peaks (number of peaks, average distance between peaks, average amplitude of peaks), rate of change (zero crossing rate) and shape (linear regression slope). This feature ex- traction process yields 42 frame-level features. As heart rate and breathing rate is only estimated during episodes of no physical movement, we apply spline interpolation to estimate missing heart rate and breathing rate values during these occasional episodes of movement. In order to find the most discriminative features for our Sleep vs. Wake classifier, we then apply a correlation-based feature selection (CFS) algorithm. Figure 5.4(a) shows a scatter plot between two top frame-level features on the activity estimate. This suggests the presence a lot of body movements during wake state and little or no movement during sleep. Figure 5.4(b) shows that heart rate tends to decrease and reaches a resting value as we transition from wakefulness to sleep.



(a) (b)

Figure 5.4: (a) Scatter plot between two frame level features: median and standard deviation of RMS energy of the low-passed (fc = 3Hz) baseband signal. (b) Scatter plot between the first quartile of RMS energy of the low-passed baseband and average amplitude of peaks of estimated heart rate.

From the feature subset selected by our CFS algorithm for different combinations of low-level features (movement, heart rate and breathing rate estimates), we trained different classifiers that are listed in Table 5.1. Using leave-one-subject-out cross-validation, we evaluated all these classifiers in terms of precision, recall and F-measure. From the results in Table 5.1, we can observe that Random Forest outperforms all other classifiers with 89.3% precision, 89.6% recall and 89.1% F-measure. When training the classifiers on just movement estimates (Mo) the performance decreases to 86.0% precision, 86.5% recall and 86.2 % F-measure. This clearly indicates that vital signal estimation (both breathing rate, Br and heart rate, Hr) carries discriminative features that enhance the performance of the classifier. Lastly the comparison between performance of two Random Forest classifiers: one trained on just vital estimates (Br and Hr) and the other on just movement (Mo) estimates, suggests that movement estimation (Mo) plays a more significant role in this Sleep vs. Wake classification.

| Features | Classifier | P(%) | R(%) | F(%) |
|---|---|---|---|---|
| Mo+Br+Hr | Naive Bayes | 82.3 | 83.5 | 81.3 |
| Mo+Br+Hr | Logistic Regression | 83.1 | 84.2 | 82.3 |
| Mo+Br+Hr | SVM | 83.1 | 84.2 | 82.3 |
| Mo+Br+Hr | Random Forest | 89.3 | 89.6 | 89.1 |
| Mo | Random Forest | 86.0 | 86.5 | 86.2 |
| Br+Hr | Random Forest | 77.8 | 80.4 | 75.9 |

Table 5.1: Sleep vs. Wake classification performance with different classifiers with different sets of features selected by CFS feature selection from different combination of low-level feature sets consisting radar-based movement (Mo), breathing rate (Br) and heart rate (Hr) estimates in terms of precision (P), recall (R) and f-measure (F).

### *5.6.2 Objective Sleep Quality Measures*

Summarizing a sleep session using a validated set of objective sleep quality measures can intuitively inform users of their sleep quality and aid in taking corrective measures such as improving sleep hygiene if needed. It also facilitates establishment of long-term trends in sleep quality. In this study we used 4 well-established sleep quality measures to summarize a sleep session. Figure 5.5 shows how sleep quality parameters are estimated from DoppleSleep's sleep vs. wake inference. (i) Sleep Onset Latency (SOL) is defined as the time taken to transition from being fully awake to being asleep. Abnormally large SOL values indicate insomnia and small SOL values indicate sleep deprivation. From figure 5.5, we can infer that the participant took about an hour to transition from wake to sleep. (ii) No. of Awakenings (NAwk) indicates the total number of transitions from sleep to wake in a particular sleep session [18]. It is another significant sleep metric that is linked with sleep apnea, disruption in circadian cycle, insomnia etc. Our participant had 3 arousal events: two long awake segments and a third brief event. (iii) Total Sleep Time (TST) is a measure of the total time duration that someone spends in sleep state in a particular sleep session. This equals to the length of sleep session minus wake time. In figure 13, the total wake time (284 mins) was subtracted from the length of sleep session (400 mins) resulting in a TST of 284 minutes. (iv) Sleep Efficiency (SE) is the ratio between total sleep time (TST) and total length of sleep session. It summarizes the three previous metrics. Reduced Sleep efficiency (below 85%) is indicative of sleep disorders since initiating (high SOL), and maintaining sleep (high NAwak and low TST) tend to be difficult. Figure 5.5 shows that our participant's TST was 284 minutes and total bed time was about 400 minutes resulting in sleep efficiency (SE) of 71%. A Pearson correlation coefficient analysis between the SOL, NAwk and SE estimated by DoppleSleep and by the ground truth reveals relatively high correlation (0.83, 0.69 and 0.78).

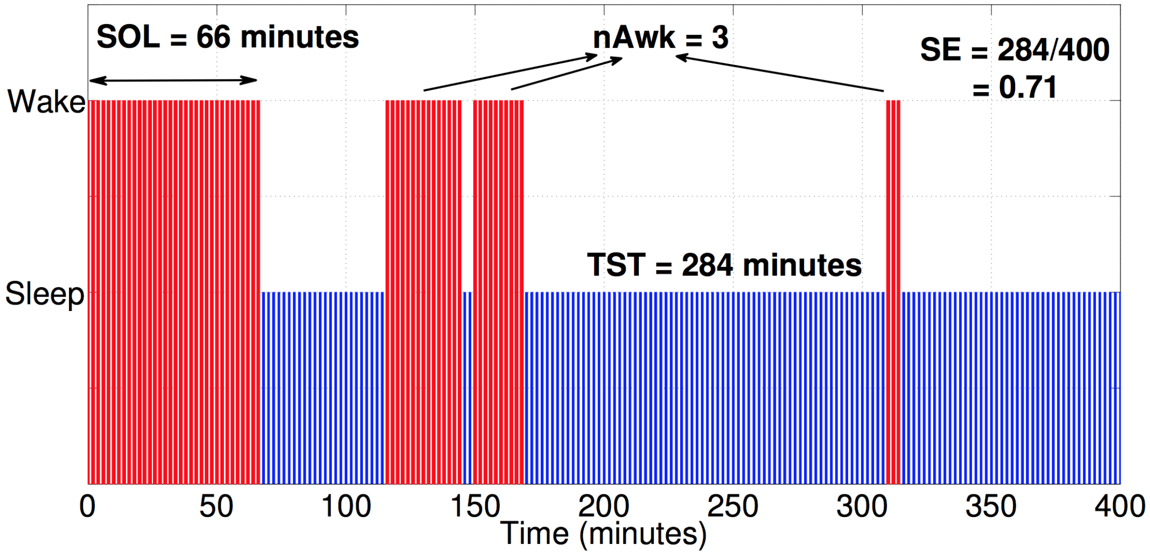


Figure 5.5: Illustration of objective sleep quality estimation

## *5.7 Summary*

In this chapter, I have described the design, implementation, and evaluation of DoppleSleep - a contactless sleep sensing system that facilitates continuous and unobtrusive long-term sleep monitoring using a single Doppler radar sensor. DoppleSleep tracks an individual's physical body movements, heart beat and breathing during sleep, and objectively infers of sleep quality. We have validated the feasibility of DoppleSleep in a lab setting and in the wild and our results show that DoppleSleep can detect physical movements with 86% recall rate, and estimate heart and breathing rate with an error rate of 8.07% and 10.84% respectively. Based on our results, by combining vital signal estimation with movement tracking, DoppleSleep shows great promise for continuous, passive, unobtrusive sleep monitoring in real-world settings with 89.6% recall for Sleep vs. Wake and 80.2% recall for REM vs. Non-REM classification.

## Chapter 6

# PROVIDING ACTIONABLE SLEEP QUALITY FEEDBACK: SLEEPAPP

### *6.1 Introduction*

In this chapter, I outline the framework for providing actionable sleep feedback through a smartphone based application SleepApp. Most commercial sleep sensors typically rely on population-level data and focus on recommendations based on objective metrics such as sleep duration or sleep efficiency. However, there is inter-individual trait-variability to sleep and people's sleep habits are individualized. To prompt users to adopt habits that improve sleep health, meaningful sleep feedback must not only provide evidence of how users' behaviors affect their sleep quality, as objectified by some of the metrics, but also show how carry-over effects of sleep affect daytime cognitive function. In this chapter, we[1] propose and validate an approach that combines both subjective and objective measures of sleep, accounting for a person's lifestyle and ties it to meaningful and measurable carryover effects such as daytime alertness and working memory. Our approach is based on the medical community's Ru-SATED framework, which characterizes sleep through six dimensions: Regularity, Satisfaction, Alertness, Timing, Efficiency and Duration. Using data collected by a smart phone app: SleepApp, with a suite of ecological momentary assessment tests from 9 participants over 14 days, we demonstrate how sleep health can be contextualized to the individual lifestyle and actionable feedback can be generated. In a follow up survey with 57 respondents, we show how the actionable feedback generated by SleepApp can encourage in users the intent to make adjustments to their sleep habits that may impact their daytime cognitive function.

### *6.2 Motivation*

It is becoming increasingly clear of the importance of sleep and its impact on our daily lives. Most people have heard recommendations from the medical community such as "get 8 hours of sleep per night" or "avoid caffeine after 2 PM." [115] However, many of these recommendations are based

---

[1] This study was conducted in collaboration with the following co-authors: Shwetak N. Patel, Julie A. Kientz

on population-level data, and individuals exhibit trait variability in sleep patterns and needs. Most research in sleep medicine has focused on either basic science about the mechanics of sleep or the identification and treatment of sleep disorders. There has been much less attention paid to providing actionable, useful feedback to the general population on how better sleep hygiene can improve quality of life from optimized sleep.

Behavior change around sleep can be difficult enough on its own, and when generic, population-level recommendations are followed and people cannot see the difference changes to sleep habits are making in their quality of life, it can be be frustrating and discouraging. In addition, most approaches to using technology for improving sleep health have focused solely on objective measures, such as sleep duration and sleep efficiency. The medical community, though it values objective data, also believes that how a person feels can be just as important as those objective numbers [58]. A person sleeping 8 hours per day consistently who does not subjectively feel well-rested is considered to still have room for improvement. In addition, knowing how aspects of sleep have a direct impact on daytime activities can be valuable knowledge for people to make changes.

We believe that a good goal for sleep technology approaches is to make personalized recommendations to individuals based on both subjective and objective data, provide actionable insights, and allow for self-experimentation. However, to reach this goal is complicated. Sleep health as a recommendation even for the general population does not have well-defined measures. Buysse et al. [26] discuss the lack of a clinical definition of sleep health even in textbooks such as Principles and Practice of Sleep Medicine [77], which mentions “sleep health” twice but does not define it. Therefore, more work is needed to understand how sleep quality feedback can both be informed by and contribute to updated clinical definitions of sleep health. In addition, providing feedback to people must be done in a way that relates their hard work in prioritizing sleep to outcomes they care about, such as alertness, cognitive functioning, or mood. If we can provide timely, contextual, and actionable feedback that shows how small changes can have an impact on their day-to-day lives, we can be more effective at promoting behavior change [80].

We aim to move sleep technology research toward the idealized goal of providing personalized, actionable feedback to people on their health based on both subjective and objective data and is grounded in holistic approaches supported by the sleep community. To accomplish this, we have developed and validated an approach to analyzing personal sleep health. Our approach is based on the medical community's Ru-SATED framework, which characterizes sleep through both subjective and objective data along six dimensions: Regularity, Satisfaction, Alertness, Timing, Efficiency, and Duration [26]. We used a smartphone app to collect data from 9 participants for 14 days along the constructs of the RuSATED framework. We designed SleepApp to provide a suite of tests for Ecological Momentary Assessment, a journal to log activities that may affect sleep quality, and a self-reported sleep log. An analysis of our data-set demonstrates how sleep health can be contextualized to the individual and actionable feedback can be generated to give people an overall ideal how small adjustments to their sleep habits may or may not impact their cognitive function, alertness, and mood. Using sample recommendations generated by SleepApp, we then conducted a survey with 57 respondents to identify if the actionable feedback would improve users' likeliness to make changes to their sleep-habits.

This research contributes a new methodological approach to adapting an in-the-wild implementation of a validated clinical framework so that it can be used understanding a person's individualized factors on sleep health. In addition, our data empirically confirms several constructs and shows how this approach can be a promising, holistic measure of sleep health that is personally contextualized and can serve as a basis for meaningful self-experimentation.

## *6.3 Study Design*

We designed our study based on the framework for sleep health proposed by Buysee et all in [26] as "a multidimensional pattern of sleep-wakefulness, adapted to individual, social, and environmental demands, that promotes physical and mental well-being. Good sleep health is characterized by subjective satisfaction, appropriate timing, adequate duration, high efficiency, and sustained alertness during waking hours." The purpose of the various tests in the study was to measure these constructs during both sleep and waking hours.

### *6.3.1 Participant Recruitment and Study Protocol*

The study consisted of 14 nights' use of the Actiwatch and 14 days of EMA from SleepApp. This length of time was chosen to mirror the two-week period used to collect baseline data in insomnia treatment [24] as well as previous work on sleep self-tracking [65]. 10 participants were selected after two rounds of screening to take part in the study. We recruited participants through various mailing lists and posting on ITHS website. From the large pool of interested and potential participants, we performed a preliminary screening through an online survey to filter participants who used an Android phone as their primary device (running a version no earlier than 4.0), who were willing to install the SleepApp software on their phone and take EMA tests at various intervals during the day, and were willing to wear an Actiwatch every night for 14 days. We received about 219 completed responses from the first survey. Since the study would take place over a long period of time and require active participation throughout the study period, we performed a second round of screening to ensure the potential participants have sufficient motivation to ensure compliance with the study protocol. In the second round of screening, participants were asked what was their motivation to take part in the study, what their usual pre-bedtime routines were and how willing they were to make changes to habits that may be affecting their sleep quality using a Likert scale. We received about 31 completed responses from the second round. From this pool, participants were chosen from different age-groups (18-75) and varying sleep schedules. During the initial study setup session, the study-coordinator met with the participants and handed them an Actiwatch and had the SleepApp software was installed on their phones. They were given a walk-through of each of the tests they were required to take on their smart phones and instructions on the use of the Actiwatch. We also received a signed consent during this session and let them know that participants were to be incentivized based on the number of consecutive nights the Actiwatch use number of consecutive EMA responses submitted. The data collected through SleepApp was uploaded to an online database (Firebase). An automatic script checked the database periodically for data entries and sent a text to the participants in case there was missing entry before the end of an EMA window. A mid-study meeting with the study coordinator was conducted to download data from the Actiwatch and check for any issues with the data an replacement Actiwatch was provided in case there were battery reset issues. At the end of the study a third and final interview was conducted to collect the Actiwatches and to provide the study compensation. Participants also filled out an end of study survey after the final interview. Nine out the ten chosen participants completed at least 10 full days of data collection for the study. One participant chose to discontinue from the study after 4 days and their data was not included in the analysis.

### *6.3.2 Actigraphy*

An Actiwatch is a device used by sleep physicians to monitor user's sleep/active cycles over long periods of time. We provided participants with an Actiwatch to for the duration of the study. All participants were required to wear the Actiwatch while sleeping at night and a text message was sent at 8pm every night reminding them to do so. We encountered several challenges from the Actiwatch during the course of the study. The devices suffered from multiple battery resets, resulting in data loss over multiple day and participants removed the watch during the day or as soon as they woke up. This led to highly fragmented and unreliable data which could not effectively be used for this study.

### *6.3.3 SleepApp Data Collection Tool*

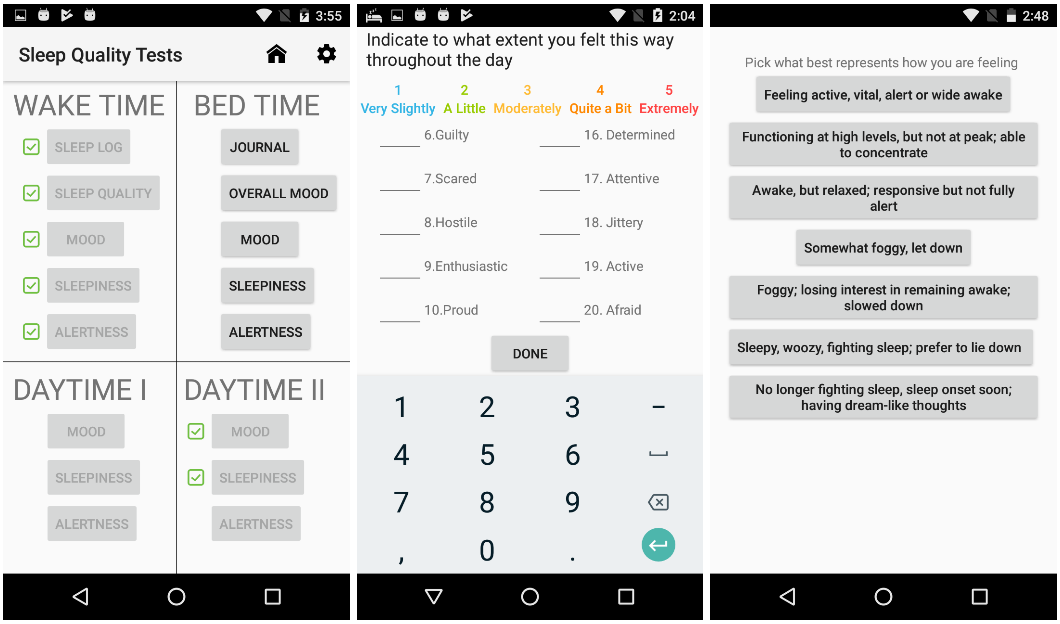
We designed an android smart phone application, SleepApp, to serve user-friendly tool to collect subjective satisfaction with sleep quality and measure the day-time carryover effects of sleep. The table below outlines the various tests administered on the smartphone through SleepApp. The app is designed as tool for Ecological Momentary Assessment (EMA) of the individual using a suite of scales combined with tests of cognitive function, mood, and alertness. We chose the tests such that each measure in the RuSATED scale was accounted for using at least one of the tests (see Table 6.1). Each of the tests has been clinically validated to be effective in measuring the intended construct.

We divided the 12-hour awake day-time period roughly into 4 windows: Wake-time to 12pm, 12pm to 4pm, 4pm to 8pm, and 8pm to Bed-time. Participants were asked for their usual wake-times and bed-time during the first deployment meeting. This frequency of EMA assessment has been used in previous work and shown to be adequate in optimizing data collection needs with minimal user burden [17]. Figure 6.1 shows sample screen-shots of the smartphone app interface that we designed for participants to perform the EMA tests. The screen was divided into four portions and the buttons leading to the corresponding scales were enabled or disabled depending on the time of day. SleepApp served pop-up notifications at the start of every window reminding participants to complete the tests for the respective window that would be available for the next 4 hours. An average of 4-5 minutes is required to complete all the EMA tests within a given window. The corresponding category of tests was enabled/disabled depending on the time-window. Once

Table 6.1: List of data sources for the study

| EMA Test/Sensor Data | Purpose | Time Administered<br>WT Wake Time<br>BT Bed Time |
|---|---|---|
| **Sleep Diary (Wake up time, sleep time and sleep quality on a scale on 1-5)** | Sleep quality self-report | WT |
| **Leeds Sleep Evaluation Questionnaire [103]** | Subjective sleep quality assessment | WT |
| **Stanford Sleepiness Scale [61]** | Sleepiness self-report | WT, 12pm, 4pm, BT |
| **PVT on smartphone [69]** | Alertness | WT, 12pm, 4pm, BT |
| **PAM (Photographic Affect Meter) [104]** | Instantaneous mood | WT, 12pm, 4pm, BT |
| **N-back test [36]** | Working memory | WT, 12pm, 4pm, BT |
| **PANAS [38]** | Mood aggregate for the day | BT |
| **Bed-time Journal** | Log of pre-bedtime activities (e.g. alcohol, caffeine consumption) | BT |
| **ActiWatch** | Movement and light data | Throughout the night |

the tests were completed, a check-mark appeared indicating completion. To encourage compliance, one hour before the end of a time-window, an email or text reminder (depending on participant preference) was sent to remind participants to complete incomplete tests if any.



Indicate to what extent you felt this way throughout the day

6.Guilty
7.Scared
8.Hostile
9.Enthusiastic
10.Proud
16. Determined
17. Attentive
18. Jittery
19. Active
20. Afraid

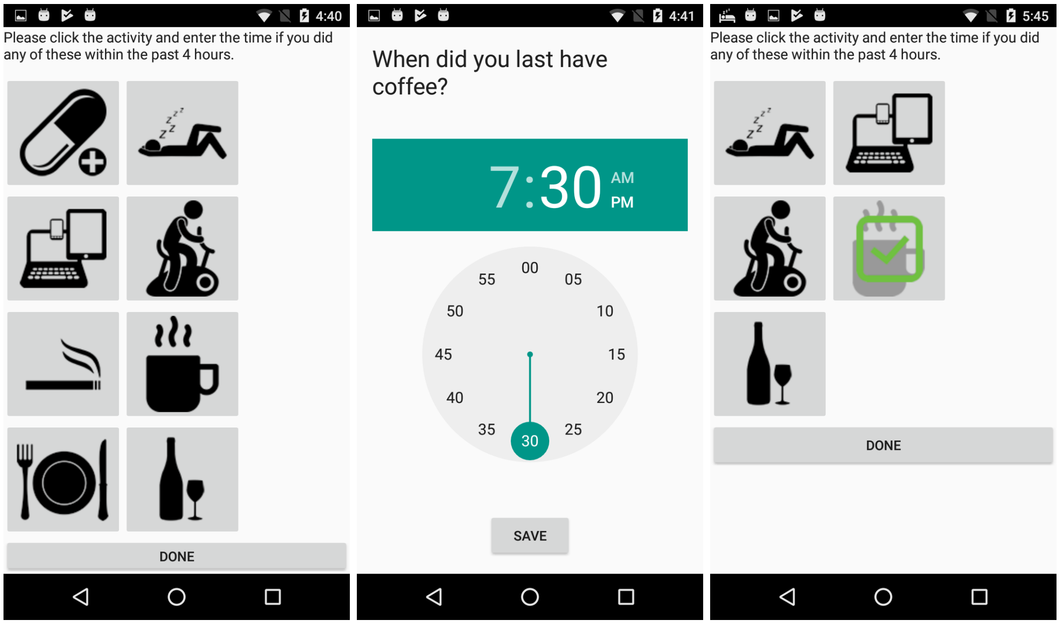
Pick what best represents how you are feeling

- Feeling active, vital, alert or wide awake
- Functioning at high levels, but not at peak; able to concentrate
- Awake, but relaxed; responsive but not fully alert
- Somewhat foggy, let down
- Foggy; losing interest in remaining awake; slowed down
- Sleepy, woozy, fighting sleep; prefer to lie down
- No longer fighting sleep, sleep onset soon; having dream-like thoughts

(a) (b) (c)

Figure 6.1: (a) Screen-shot of the different windows during which the users will be required to log self-reported data; (b) Screen-shot of the PANAS scale to log mood aggregate for the day; (c) Screen-shot of the Stanford Sleepiness Scale (SSS) on SleepApp.

In addition to the EMA tests participants were also asked to record pre-bedtime activities through a bed-time journal. Figure 6.2 shows the screen-shot of the journal designed for SleepApp. At the beginning of the study participants were allowed to chose variables that they were were probable to engage in. From each journal entry, the participants were allowed to retrospectively log the particular time at which the performed an activity if it fell within the bed-time window. The list of pre-bedtime activities that could have an effect on sleep quality was derived from clinical recommendation for sleep hygiene [63, 115].



Please click the activity and enter the time if you did any of these within the past 4 hours.

When did you last have coffee?

Please click the activity and enter the time if you did any of these within the past 4 hours.

(a) (b) (c)

Figure 6.2: (a) Screen shot of the variables available for the user to log as pre-bedtime activities. (b) Screen shot of the timer to log the time at which an activity was performed. (c) Screen shot of a customized sleep journal with a completed entry

## *6.4 Analysis*

At the end of the study, we downloaded participants' data from Firebase and performed these preliminary steps to remove outliers in the data. We constructed the following heuristics assuming that participants might make erroneous or unintentional entries in haste that they did not have the chance to rectify since the buttons were disabled after an entry was completed. Some erroneous data entries also appeared due to glitches in the app. While the app behaved in the manner for which it was designed for most of the participants, for some, there were instances where the buttons were not disabled after an entry resulting in redundant entries.

- Sleep log entries with missing fields or entries with the exact same sleep time and wake time resulting in a sleep duration of 0 hours were removed.

- We added or subtracted 12 hours to the sleeptime/waketime entry in cases where the participants might have forgotten to toggle the AM/PM switch on the app.
- If the first entry of any test fell outside the normal range of values for the participant or if it was entered during a time that was outside the intended window, it was discarded because it could have been an entry logged during the study set-up meeting.
- We disregarded all but the last entry in cases where there were redundant entries due to a bug in the app.
- We removed outlier sessions for the PVT if the mean reaction time for a session fell outside the mean ± 2.5*SD for that participant [17]
- Nback test entries where there were 0 number of total clicks for an entire session were deleted.

### *6.4.1 Dependant and Independent Variables Affecting Sleep*

The data gathered in this study can be divided into three domains as shown in Table **??**. The first column lists the variables that could affect sleep quality that was measured during the study through SleepApp. This is certainly not an exhaustive list since there are several environmental variables, psychological and physiological variables that affect sleep but measuring them was beyond the scope of this current study. The second column lists the metrics used to measure sleep quality. The independent variables listed in the first column effect the variables in the second column - the dependant variables. Previous work has shown that sleep health is not sufficiently characterized by the quality of sleep at night alone but includes day-time carryover effects as well. These day-time carry over effects of sleep such as variation in mood, alertness and cognitive function are listed in the third column. While these are dependant variables as well, they do not have a direct correlation with the independent variables in the first column. They are an extension of the dependant variables in the second column. This distinction, between variables in the first, second and third column, helps us draw make pair-wise correlations and to distinguish between the combination of independent variables and measurable dependent variable outcomes in the linear regression model discussed in the results section.

| Variables that affect sleep | Sleep Metrics | Effects of Sleep |
|---|---|---|
| Timing of sleep | Satisfaction | Alertness |
| Duration of sleep | Efficiency | Mood |
| Exercise | Sleep Onset Latency | Working memory |
| Caffeine | Number of awakenings | |
| Large Meal | Restlessness | |
| Alcohol | Duration of sleep | |
| Tobacco | | |
| Medication | | |
| Electronics Usage (Blue light) | | |
| Napping | | |
| Mood | | |

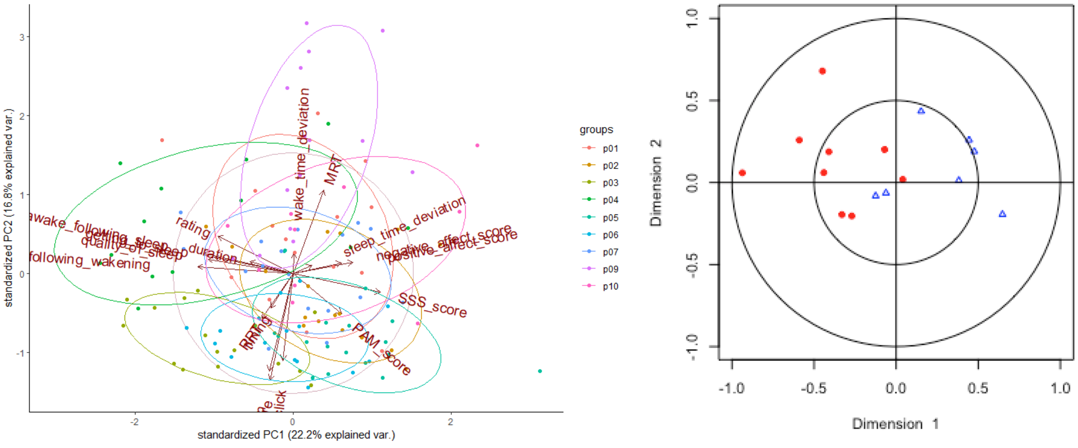
Table 6.2: List of the dependent and independent variables that affect sleep

## *6.5 Results*

### *6.5.1 Manifestation of Inter-individual Trait Variability in Sleep*

Sleep physicians have stated in past work [118] that there's inter-individual trait variability to sleep. Sleep need and the effects of sleep deprivation can vary from person to person. To test if this is reflected in our data, we performed a Principal Component Analysis. PCA is a technique used to emphasize variation and bring out strong patterns in a dataset using orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. This transformation is defined such that the first principal component has the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. Therefore, using PCA, we can transform the data to highlight the variation in the data across the different variables.

Figure 6.3(a) shows the PCA of the data collected from the 9 participants. The ellipses in the plot correspond to individual participants' data. This shows there is strong participant-dependant variation in the results from the tests as exhibited by the clustering of the ellipses along the corresponding variable axes. This emphasizes the inter-individual trait variability in sleep quality and its effects. Example, P09 is characterized by high values for Mean Reaction Time and Wake-Time deviation whereas P03, P04 and P05 were characterized by measures of cognitive function such as percentage of correct responses from the Nback test and subjective sleep quality rating. Even from a limited sample size, we can see the effects of these inter-individual traits.



(a)

(b)

Figure 6.3: Plot of data-set along first and second principal components across the nine participants showing variance between the dependant and independent variables

### 6.5.2 *Canonical Correlation Analysis*

In order to get a summary statistic of the dependant and independent variables in our data, we performed Canonical correlational analysis on the data. Unlike PCA, Canonical correlational analysis find the dimensions that maximizes the variation between a linear combination of dependent variables and a linear combination of independent variables. Figure 6.3(b) shows the plot of the IV (Independent Variables) and DV (Dependent Variables) along the first and second dimensions. From the spread of the variables, we can see there's enough variation across a combination dependent and independent variables.

### *6.5.3 Correlations between the Dependent and Independent Variables*

Repeated measures correlation (rmcorr) is a statistical technique for determining the common within-individual association for paired measures assessed on two or more occasions for multiple individuals. Using the rmcorr function in R, we tested the data for correlation against repeated measures (since data for each pair of correlations was collected repeatedly over 2 weeks) with each participant as the factor.

### *6.5.4 Subjective Sleep Satisfaction Correlates with Sleep Duration*

We can see a highly significant correlation ($p<0.001$) between duration of sleep against subjective sleep quality rating (see Figure 6.4). This a fairly straightforward correlation that confirms previous work [42]

### *6.5.5 Validity of the LEEDS scale*

The Leeds Sleep Evaluation Questionnaire comprises ten self-rating 100-mm-line analogue questions concerned with aspects of sleep and early morning behaviour. The questionnaire has been used to monitor subjectively perceived changes in sleep. We tested participants' answers to the LEEDS scale with the Cronbach's alpha to confirm the internal consistency of the scale. The standard alpha value was 0.81, thus showing high internal consistency.

We also tested the validity of the scale by finding the repeated measures of association between subjective sleep rating and answers to the Quality of Sleep (QOS) section of the scale and found a significant correlation ($p<0.001$). There was also a significant correlation ($p<0.001$) between the Behaviour Following Wakening section of the scale and the participants' responses to the Stanford Sleepiness Scale. This allows us to use the LEEDS scale measures to get finer details of subjective sleep sleep quality assessment compared to the 5-point sleep quality self-report from the sleep journal or the 7-point Stanford Sleepiness Scale.

### *6.5.6 Sleep-time Correlates with Ease of Getting to Sleep*

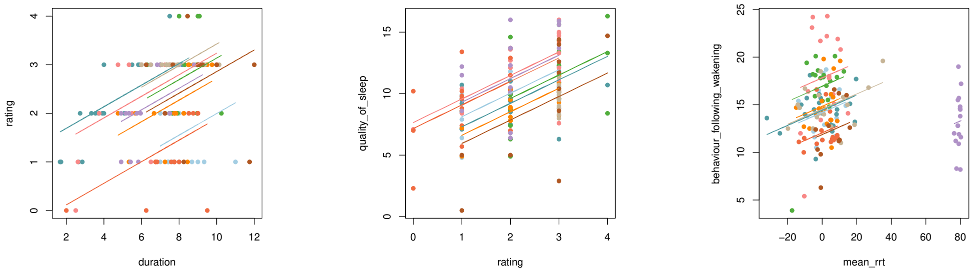
To extract a measure of regularity we calculated the sleep-time deviation for each day for each participant. From the LEEDS sleep evaluation questionnaire, the answers to the question "How would you describe the way you currently fall asleep in comparison to usual?" (More difficult - Easier than usual, Slower - More quickly than usual, Less sleepy - More sleepy than usual) was summed up to get a GTS (getting to sleep) score. From our data set, we found a highly significant negative correlation ($p < 0.001$) between sleep-time deviation and the GTS score. This shows that the closer the participants' sleep-time for the night was to the mean sleep-time, the easier and faster it was for them to fall asleep. This is in accordance with medical community guidelines for healthy sleep habits: maintaining a regular sleep-time and wake-time.

### *6.5.7 Daytime Alertness Correlates with Behavior Following Waking*

We used Relative Response Time from the PVT as a measure of Alertness in the study. This method has shown to be effective in previous work [17, 70]. There was a significant positive correlation ($p< 0.01$) between the mean RRT for a participant over the course of an entire day, and the self-reported behaviour following awakening. Responses to the BFW section of the scale included how tired or alert they felt when they woke up and their balance and coordination upon awakening. The more alert participants felt when they woke up, the higher their relative alertness was throughout the day. This shows there is a correlation between behaviour following wakening and overall alertness throughout the day, further highlighting the extended day-time carry-over effects of sleep quality.

### *6.5.8 Daytime Alertness Correlates with Sleep Duration*

The effect of sleep deprivation on alertness has been established in the medical community through several studies [17]. We were able to replicate the findings in our study as seen by the significant correlation ($p = 0.01$) between mean relative reaction time and sleep duration.



(a) a (b) b (c) c

Figure 6.4: (a) Plot of the repeated measures correlation between (a)duration and rating (b)sleep-time deviation and LEEDS scale measure Getting to Sleep (c) Subjective sleep quality rating and LEEDS scale measure Quality of Sleep

### *6.5.9 Recommendations from the Linear Model*

Previous work [80] has shown that people can be led toward behaviour change through SMART (Specific, Measurable, Actionable, Realistic, and Timely) goals. As a start, from our data set, we built a simple regression model by combining the independent variables to generate outcomes for the dependant variables. By affecting changes to the independent variables through their pre-bedtime activities and by following the guidelines for sleep hygiene, participants will be able to see positive outcomes in terms of increased day-time alertness and cognitive function.

### *6.5.10 Independent Variables Affecting Alertness*

In the first formula, we included all the independent variables that affect sleep quality to get an idea about which of the sleep metrics have an effect on day-time alertness.

$$lm(formula = mean(rrt) \sim duration + rating + timing + sleep - timedeviation + wake - timedeviation + \cap$$
$$gettingtosleep + qualityofsleep + awakefollowingsleep + behaviourfollowingwakening$$
$$\cap + alcohol + meal + exercise + electronics, data = all - combined) \tag{6.1}$$

The following variables had a significant effect on the relative mean day-time alertness measured using the PVT listed in the order of increasing significance : timing of sleep (Pr(>|t|) = 0.05), quality of sleep (Pr(>|t|) <0.05), sleep time deviation (Pr(>|t|) = 0.01), behaviour following wakening (Pr(>|t|) <= 0.001), meal(Pr(>|t|) <0.001) and electronics usage (Pr(>|t|) <0.001).

The relatively high significance of the the two independent behavioural variables (electronics usage and having a large meal before sleep) may be attributed to that fact that these are participant specific effects: these variables may not have been logged by all participants, and this simple linear regression model does not take into account the repeated measures from a single participant as a random effect. However, the results still highlights how behaviour variables have a cascading effect on the quality of sleep and in turn affect the day time carry over effects of sleep. Using the coefficients derived from this model, we can serve specific recommendations to users that ties behavioural changes to measurable effects in the outcome, increased or decreased alertness the next day.

### *6.5.11 Actionable Variables that Affect Alertness*

To see the effect of the only the actionable variables on day-time alertness, we proceeded to remove all of the independent variables from the model that participants do not have direct control over (such as subjective sleep quality assessments).

Similar to the results from the first model, there was a significant correlation between sleep-time deviation (Pr(>|t|)<0.05), meal (Pr(>|t|)<0.001) and electronics usage(Pr(>|t|)<0.001). Avoiding consuming large meals before bed and exposing oneself to blue light from electronic devices are in-line with the sleep medicine community's recommendation for good sleep hygiene.

### *6.5.12 Actionable Variables that Affect Working Memory*

In our study, we used the nback test (n = 2) to measure changes in working memory in our study. We used percentage of right rejections (not clicking on a shape that was similar to the one displayed two shapes ago) and percentage of right hits (clicking on shape which is similar to the one displayed two shapes ago) as a measure to calculate performance on the nback test. Although we generated an aggregate score by combining the number of right hits and right rejections, we found that the aggregate score did not sufficiently capture the variance in participants' performance since the percentage of right rejections artificially boosted the overall score. In the linear regression model below, we combined the actionable independent variables to see the effect on mean percentage right clicks per day.

$$\begin{gathered} lm(formula = mean - percentage - right - clicks \sim duration + timing + sleep - time - deviation \cap \\ + wake - time - deviation + alcohol + meal + exercise + electronics, data = all - combined) \end{gathered} \tag{6.2}$$

The following variables showed a correlation to working memory in increasing order of significance: sleep duration (Pr(>|t|) <0.05), sleep time deviation (Pr(>|t|) <0.05) and consumption of alcohol (Pr(>|t|)<0.001). The consumption of alcohol is known to cause disruptions in sleep cycles, therefore it's effects can be seen through decline in cognitive function the following day. Here again, the relatively high significance can be attributed to the model not accounting for participant specific effects.

Next, we tested to see which variables had an effect on percentage of right rejections. This score was relatively high across all the participants since non-action (not clicking on any shape at all) would still give a 100% right rejection rate.

$$\begin{gathered} lm(formula = mean - per - right - reject\ duration + timing + sleep - timedeviation \cap \\ + wake - time - deviation + alcohol + meal + exercise + electronics, data = all_combined) \end{gathered} \tag{6.3}$$

We can see a high significance for exercising in the pre-bedtime window (Pr(>|t|) = 0.05), alcohol consumption (Pr(>|t| < 0.05), and electronics usage before bed (Pr(>|t|)<0.01) with respect to percentage of right rejections in the Nback test.

### *6.5.13 Generating Specific and Actionable Feedback*

In addition to the looking at the p values, we use the coefficients generated from the linear models presented above to provide specific and measurable feedback. For example, in the first model for day-time alertness, duration had a -1.675 estimate with 0.68 standard error. Because the duration of sleep was calculated in hours, the model shows that changing the duration by 1.6 hours can have a significant effect on day-time alertness. The generalized model built from the data across the nine participants can also be used to serve personalized recommendation by looking at how an individual's data deviates from the mean of the population. If an individual's data is several standard deviations away, the can be asked to make changes to that specific variable in order to affect changes in the outcome.

### *6.5.14 Participant Feedback on SleepApp Data Collection Tool*

We conducted a post-study survey to determine the user-burden of the SleepApp data collection tool. Eight out of nine participants felt SleepApp demanded too much mental effort a little bit of the time or sometime. Six participants said SleepApp was not at all difficult to use, while three said SleepApp was a little bit or somewhat difficult to use. Six participants said the frequency of tests was not at all high or a little bit high. Two participants said it was very high and one said extremely high. However, seven participants said the number of tests within each window was not at all high and 2 said somewhat high. Five participants stated that the Alertness tests demanded too much mental effort and took a long time to do. Most of the other EMA tests did not demand much mental effort at all according to the survey responses. Six participants responded that remembering to use SleepApp was hard.

## *6.6 Survey on Actionable Feedback Prompted by SleepApp*

To find out if the recommendations generated by SleeApp would be useful and would promote behaviour changes, we conducted a survey with 57 respondents on Mturk. We asked participants to respond to the PSQI questionnaire to determine their overall sleep health to see if the sleep health distribution of the respondents was similar to that of a normal population. The survey was divided into two main sections. The first portion of the survey was structured around participants sleep hygiene and their awareness about how these environmental and behavioral variables affected their sleep quality and by extension, their day-time functioning. The second part of the survey consisted of example scenarios of recommendations generated by SleepApp and the participants likeliness to make behaviour changes based on the scenarios. Specifically we set out to answer the follow questions through the survey:

- What are the most common self-reported behavioral variables that affect sleep quality?
- How aware are participants of the effects of these behaviours on their sleep quality without doing a controlled self-experimentation study?
- How likely are respondents to avoid these behaviors that affect sleep if it affected their daytime cognitive function?
- Do respondents likeliness to make behavior changes correspond to different levels of improvement or decrement in cognitive function?

Answers to these questions can provide us with key insights into the actual design of SleepApp recommendations. Since the design of actual recommendations and generating the outcomes are beyond the scope of this preliminary exploration, survey respondents were given hypothetical scenarios with regard to feedback generated by SleepApp and asked respond based on those scenarios.

### *6.6.1 Questions on Sleep Hygiene*

In order to figure out the distribution of factors affecting sleep quality among respondents and their relevance to the users' lifestyle, they were asked to select from a list, all the factors that they thought affected their sleep quality such as regularity in their sleep schedule, napping during the day, caffeine, alcohol and nicotine intake, electronics usage and engaging in high intensity exercise before bedtime. Respondents were also asked about environmental variables that affect sleep quality although those factors were not included in the current SleepApp prediction model.

### 6.6.2 *Questions on Self-experimentation*

People generally have a good subjective assessment of their overall sleep quality and how their behaviours affect their sleep quality. The goal of SleepApp is to supplement those subjective assessments with objective data gathered from the validated tests. In order to get a sense of how people created their own mental models of the effect of their behaviours on sleep quality, there were asked to choose from a list, the ways in which their sleep quality was affected (the dependant variables) such as increased amount of time taken to fall asleep, increased restlessness, increased number of wake-times during sleep, reduced satisfaction and increased duration of sleep. Respondents were also asked to identify how their daytime cognitive function was affected by their sleep quality. In addition to the variables that were included in SleepApp's prediction model such as working memory, alertness, changes in mood, and drowsiness during the day, variables such as ability to focus were also included in the list. Respondents were also asked about the amount of time it took for them to notice patterns in the relationship between sleep the dependant and independent variables, starting from a several days to several weeks, and if they had used any device or a journal to observe patterns in their sleep quality.

### 6.6.3 *Questions on Behavior Change Prompted by SleepApp*

They survey was designed to see if the recommendations from the sleep model developed by SleepApp would be actionable enough that users might make changes to their behaviour in order to improve their daytime cognitive function. To this end, respondents were initially asked how likely they were to make changes to behaviors for an improvement in cognitive function the following day. We wanted to see if respondents were more likely to certain behaviour changes but not others. Following that respondents were given four different scenarios and asked how likely they would be to make the following behavior changes, if prompted by SleepApp, for a specific percentage of increase or decrease in cognitive function:

- I would try to go to bed and wake up at around the same time everyday, If SleepApp predicted that my alertness would improve by ___
- I would try to avoid screen-time (exposure to blue light through electronics usage) 30 mins

before I go to bed, if SleepApp predicted my memory would improve to be ___

- I would avoid going to bed later than usual, if SleepApp predicted that my alertness would decrease by ___by staying up late
- I would try to avoid screen-time (exposure to blue light through electronics usage) 30 mins before I go to bed, if SleepApp predicted my memory would decrease by ___by engaging in screen time.

The options for the percentage increase and decrease were derived from the normal curve equivalents of a normal distribution: 50%, 70% and 90% for improvement and 10%, 30% and 50% for decrement. The respondents were presented with options on a 7 point Likert scale starting from very unlikely to very likely.

## *6.7 Survey Results*

The results of the survey highlight the significance of relevant contextual sleep feedback in promoting healthy sleep habits.

### *6.7.1 Sleep hygiene*

While a majority of the respondents practiced good sleep hygiene such as going to bed and waking up at the same time (62%) and not napping during the day for more than an hour (72%), there were still some behavioral variables that were common among respondents that were detrimental to good sleep health. About 50% of respondents did not maintain a consistent sleep schedule across all 7 days of the week and 70% admitted to engaging in screen-time (leading to exposure to blue light) 30 mins before bedtime. In addition to behavioral variables, environmental factors that affect sleep quality such as loud noises due to construction (42%) and bright lights(46%) coming into the bedroom were more commonly reported. SleepApp focuses on providing recommendations that are related to behavioral variables since making changes to environmental variables may not always be actionable.

### 6.7.2 *Likeliness to take action prompted by SleepApp recommendations*

From Figure 6.5(a,b) we can see that more than 85% of all respondents reported that they would find the recommendations from SleepApp to be useful every night, every morning and in the form of a weekly report or focusing on monthly trends. While respondents were more likely to find the nightly recommendations useful(89.5%), they were not as likely to follow the recommendations from SleepApp every night(71.9%). On the other hand, respondents were more likely to follow the recommendations on the eve of an important event such as a examination or a presentation(85.9%) even thought they were less likely to find the recommendation useful(78.9%). Figure 6.5(c) shows that although more than half of all respondents were likely to avoid behaviours that would affect their sleep quality if prompted by SleepApp, respondents were less likely to make changes to certain behaviours such as abstaining from the use of electronics before bed.

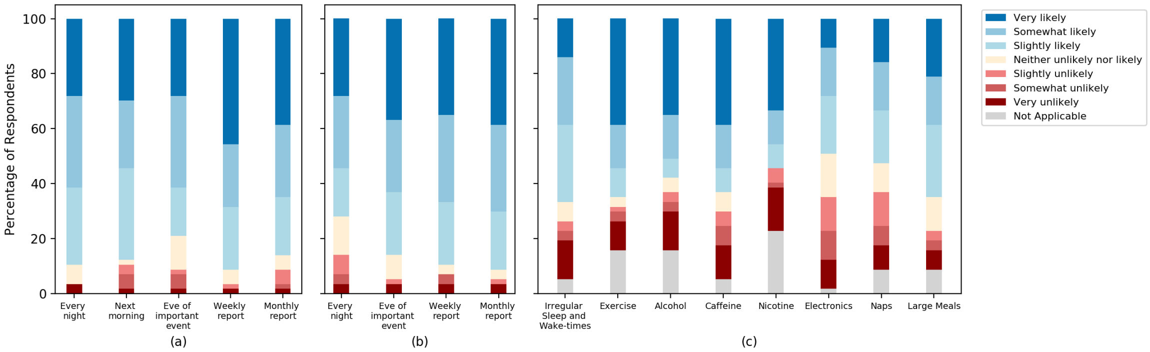


Figure 6.5: (a) Percentage of respondents who were likely to find recommendations from SleepApp to be useful. (b) Percentage of respondents who were likely to follow the recommendations from SleepApp. (c) Percentage of respondents who were likely to make changes to the factors affecting sleep quality if prompted by SleepApp

### 6.7.3 *Intent to make behaviour change dependant on daytime cognitive function*

As expected, Figure 6.6the higher the percentage of improvement in cognitive function, the more likely respondents' intent was to make behaviour changes. However, an interesting finding was that respondents were more likely to make changes to their behaviour if SleepApp predicted an improvement in cognitive function than to avoid a predicted decrease in cognitive function the following day. There was no significant difference between how likely respondents were to intent to maintain a consistent bedtime versus avoiding electronics for an improvement or to avoid a decrease in cognitive function.

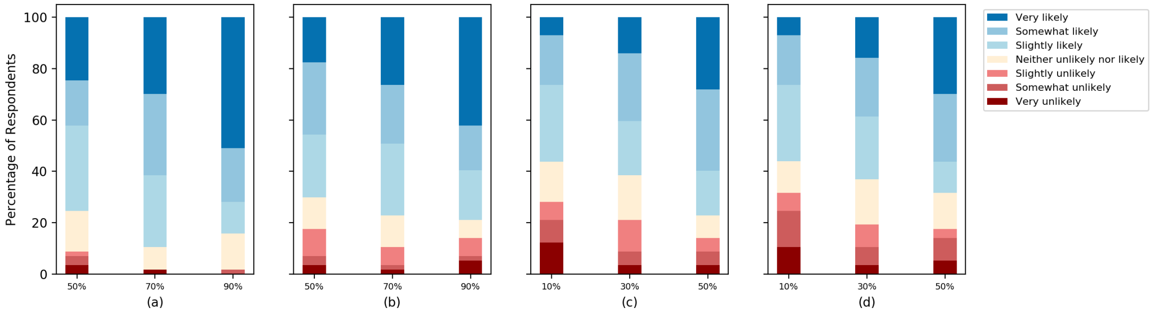


Figure 6.6: (a) Percentage of respondents who were likely to say they would make the following behaviour changes for an increase or decrease in cognitive function the following day: (a)maintain a consistent sleep schedule if SleepApp predicted improvement in alertness. (b)avoid engaging in screen time if SleepApp predicted improvement in working memory. (c)avoid going to bed 30 minutes later if SleepApp predicted reduced alertness (d) avoid engaging in screen-time if SleepApp predicted reduced working memory.

### *6.7.4 SleepApp Accuracy*

Although respondents were more likely say they would follow SleepApp's recommendation for higher levels of accuracy, Figure 6.7(a) shows that respondents' overall likeliness to follow SleepApp's recommendations remained consistent across the three levels of precision (0.66, 0.8, 1.0), meaning respondents were likely say they would follow recommendations if it matched their own subjective assessment at least 7 out of 10 times. Among the mechanisms that they would adopt to compensate during days of predicted reduced cognitive function, more than 50% of respondents reported that they would incorporate physical activity, avoid driving or operating heavy machinery and try to go to bed earlier. 90% of respondents reported that would use SleepApp if it were available to them in the next 6 months.

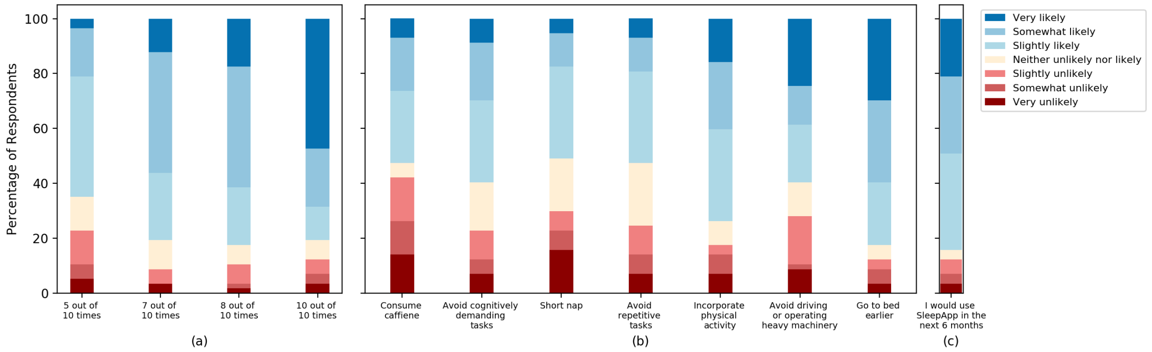


Figure 6.7: Percentage of respondents who would intend to (a)follow recommendations from SleepApp if it matched their subjective assessment (b)employ compensation mechanism on day of predicted decline in cognitive function (c)use SleepApp if available to them in the next 6 months.

## *6.8 Insights from the survey*

The findings from the survey offer some key insights into how the recommendations from SleepApp can be made actionable.

- Since electronics usage is a sleep hygiene factor that affects a majority of the users, monitoring smartphone usage passively on the app and generating reminders before bedtime could nudge users' towards decreased usage before bedtime.
- Giving users a choice of recommendations based on their preference of variables they are willing to change, could increase chances of people making changes to those particular variables.
- Presenting feedback in terms of improvement in cognitive function rather than decrease in cognitive function may facilitate likeliness to make changes to pre-bedtime routines. Since the survey data suggests that a prediction of reduced cognitive function is not necessarily a deterrent to behavior that negatively affect sleep quality.

## *6.9 Discussion*

Our findings show that sleep health can be defined by not merely relying on objective sleep metrics, but by incorporating subjective assessment and by correlating behavioral variables from a person's lifestyle that may affect sleep quality as well as the daytime carryover effects of sleep on mood, cognitive functioning, and alertness. These specific, measurable, and actionable goals from the study seem promising in that they have the potential to help people understand their sleep in the context of their lifestyle. However, there is a need to further investigate the best way to deliver this feedback to the user.

### *6.9.1 Facilitating Self-Experimentation*

The findings from our study serve as a baseline, allowing users to explore variables that affect sleep without controlling for any specific variables or factors. However, to confirm the presence of a correlation between an independent variable and a corresponding outcome. While the data collected for this study can serve as a preliminary hypothesis generation phase, people will still need to carry out self-experiments to test their hypothesis. For example, to test if alcohol consumption does have the suggested effect, they would have to perform a trial that controls for confounds, such as an single case study ABAB design [39]. The guidelines for carrying out such self-experiments have been discussed in TummyTrials [64] and in Daskalova et.al's work [40]. Unlike with Irritable Bowel Syndrome, the use case in TummyTrials, people may not have full control over the independent variables they are trying to manipulate. For example, while maintaining a regular sleeptime schedule may seem doable, one cannot physically control for the amount of hours spent in sleep (i.e., people cannot forcibly sleep for a certain duration). While these confounds can not be entirely avoided in designing self-experiments, we can account for this by allowing for a certain amount for uncertainty in the feedback given to users. Also, as shown in previous work [64], allowing flexibility in the design of self-experiments and the option of allowing users to customize variables can help make self-experiments successful. Specifically, providing feedback by ranking the independent variables in the order of significance or strength of correlation can help people choose which variables they would like to manipulate in their self-experiments. In some cases, it might not be feasible for people to control for the variable due to unforeseen circumstances. Allowing for flexibility in the self-experiment design can motivate people to continue on their trial without having to sacrifice several days worth of data due to a few missing data-points.

### *6.9.2 Timing of Feedback*

An important aspect of delivering health feedback to people tracking their data is the timing. Specifically, at what points in time they are allowed to see the findings from their data. For example, allowing people to see objective measures collected from the tests or sensors might alter their subjective perception of the measured construct. On the contrary, not providing timely feedback may lead the users to disengage from continued tracking. Therefore we need to be wary about how and when people get to see the feedback. Nahum-Shani et al have proposed a framework in [98] for delivering just-in-time interventions, while adapting to an individual's changing circumstance. The model proposed in the paper can serve as a concrete guideline for delivering timely feedback.

## ***6.10 Summary***

Sleep technology approaches that can make personalized recommendations to individuals based on both subjective and objective data and provide actionable insights and allow for self-experimentation can help people take measures to improve sleep health. Current commercial sleep tracking systems focus largely on objective measures and fall short in helping people make meaningful connections between sleep quality and its day-time carry over effects. In this chapter, we have proposed an approach to adapting an in-the-wild implementation of a validated clinical framework that can be used for long-term tracking. In addition, our data empirically confirms several constructs and shows how this approach can be a promising, holistic measure of sleep health that is personally contextualized and can serve as a basis for meaningful self-experimentation. We have also discussed the limitations of the current approach and how they can be overcome in the future.

# Chapter 7

# LIMITATIONS AND OPPORTUNITIES FOR FUTURE WORK

## *7.1 Introduction*

In this chapter I have discussed limitations of the studies I have presented in this thesis work. I have proposed opportunities for future work that can overcome these limitations. Finally, I have discussed meta-lessons learned I have learned during the course of my research.

### *7.1.1 Limitations in WiBreathe and DoppleSleep*

Through the design and implementation of WiBreathe and DoppleSleep, I have shown the possibility of providing a non-contact and reliable long-term sleep tracking solution. However, there are limitations in the system that can be overcome through further research in order to implement the system in users' home on a commercial level. The 2.4GHz wireless signal used in the WiBreathe system was produced on a custom FPGA provided through an SDR(software defined radio) and required highly directional antennas. This is not a scalable option since the hardware costs thousands of dollars and is not easily available to users. While DoppleSleep provides a low-cost, off-the shelf option for wireless sensing, the continuous wave (CW) Doppler RADAR is limited in its capability to distinguish physiological signals between multiple users. Since a large percentage of users sleep with a partner, being able to distinguish signals from multiple users is critical to the real world implementation of these systems.

### *7.1.2 Limitations in SleepApp*

From the preliminary results shown through the SleepApp study, we can see that although there is high significance in the observed correlations, the strength of the correlations were relatively small. This could be strengthened by collecting data for a longer duration (n  40 days) from a larger pool of participants. Having a larger data-set can also increase the variance in the data allowing for better recommendations.

It is well known that the variables affecting sleep quality are not purely independent and that there may be several confounds between variables affecting sleep. This applied to the data from our study as well. For example, there are several environmental variables such as ambient light, sound, or co-sleeping with partner as well as physiological variables such as the onset of a sickness that may affect sleep quality that are not accounted for in the linear model used to predict day-time outcomes. Self-experiments with randomized conditions may help with eliminating some of these confounds.

While commercial devices are not subject to the same validation standards as those of a clinical device, using commercial sleep tracking devices could have given us better access to the data. The Actiwatch was limited in its functionality in several ways. There was no interface available in the hardware to inform participants or the study-coordinator of any possible malfunction. There was also no option of periodically uploading the data to an online server which could have allowed us to periodically check for data quality and compliance. Using a commercial sleep sensing device might have been helped us overcome these issues and they are relatively inexpensive when compared to the Actiwatch.

## *7.2 Future Work*

While I have tried to take a holistic approach to sleep tracking through my research, there are still several avenues that are open to be exploration and advancements that can be made to enable users with with technology and tools to improve their sleep health and general well-being.

### *7.2.1 Innovation In Sleep Sensing*

Sensor development for sleep tracking comes with an inherent necessity to strike a delicate balance between comfort and accuracy. Since sleep plays a critical role in people's lives, disturbances in their sleep environment from intrusive sensors (bright LEDs etc.), concerns about invasion of privacy (audio recordings) and cumbersome wearables can discourage people from engaging in long-term sleep tracking. The issues regarding privacy and comfort are especially relevant since the goal of at-home sleep-tracking is long-term, and not restricted to a single night or a few weeks, as is the case in clinical tracking. While the more accurate sensing devices tend to be more cumbersome and the lesser intrusive devices not as accurate and sensitive, recent developments in remote sensing technology have minimized this trade-off between accuracy. In my research, I have demonstrated how a low-cost, easily deployable 24 GHz Doppler RADAR can enable long-term sleep tracking while minimally invading the users' sleep environment.

### *7.2.2 Sleep tracking for more than one user*

Reliable sleep tracking for more than one user is a critical step in non-contact sensing using wireless signals. A huge percentage of the population sleep with partners and being able to (i) effectively sense sleep of a target user in the presence of another person or (ii) to track sleep sleep patterns of both users simultaneously is critical to the successful implementation of wireless sensing in the wild. This can be done using the following methods.

### *7.2.3 Using Alternative RADAR Modulation Schemes*

There is promise in using FMCW RADAR systems for multiple person detection since FMCW signals can transpose separation in the spatial dimension to separation in the frequency domain. In the ideal scenario if two users are separated by a distance farther enough to be resolved by the spatial resolution of the signal, their physiological signals can be effectively captured by different frequency channels. One caveat of using this technique is that while the multiple physiological signals can be effectively captured, they cannot be identified or assigned to the respective user, without prior knowledge of the user's relative position to the signal. While it maybe safe to assume that user's positions may change but will not completely switch, during the course of the night, this is a drawback of this particular sensing modality, something that will not be an issue for wearable sensing methods. On approach to solve this issue is by learning long-term trends in a user's sleep pattern, machine learning models can identify and resolve signals from multiple users sharing the same space.

### *7.2.4 Source Separation Algorithms*

Wireless signals are subject to attenuation and carry noise due to several factors. The signal received at the receiver is subject to path loss and also carries noise added by the objects placed in the surrounding environment. While WiBreathe and DoppleSleep, resolve these by measure such as filtering and by using the adaptive algorithm there are more sophisticated algorithms that can be implemented.

### *Using Commodity Wi-Fi Devices*

There have been several breakthroughs in the scientific community in the past few years using WiFi signals from commodity devices for fine-grained indoor localization [76] as well as gesture recognition [105]. These studies leveraged the underlying CSI and RSSI information from commodity WiFi chipsets in laptops to provide centimeter level resolution that can be used to identify human gestures and movements. Estimating Angle of Arrival (AoA) [107] and Time of Flight (ToF) [78] at the receiver using multiple antennas (or an antenna array) has been a widely used approach in the field of localization and communication using wireless signals. Multiple antennas placed at distance intervals that are a function of the wavelength of the carrier frequency, receive time-delayed copies of the transmitted signal can be used in interesting ways. In situations where the received signal is a composite of several reflected signals due to multipath propagation, AOA can be used for source separation of the various multi-path components. Using algorithms such as MUSIC [22] and ESPIRIT [94], the AoA of different signal sources, such as vital signals from multiple people in the room can be deconstructed and measured effectively.

The advantage of using commodity WiFi devices is manifold for this particular use-case (i) multiple access points for WiFi are commonly available in indoor spaces these days, eliminating the need for additional hardware. (ii) as the number of antennas determines the number of multipath components that can be resolved, multiple sub-carriers (up to 30) of the OFDM modulation scheme of commodity WIFi, along with the in-built 3-antenna array available in routers can be leveraged to increase the resolution of the system to that of a 30*3 = 90 sensor-array system [76]. While these algorithms can enhance the resolution of the system to a centimeter level, the fundamental physical limitations of signal in-terms of the maximum allowable distance between the two target sources needs to be determined. Enabling vital-signal detection of multiple persons through commodity WiFi routers, can be beneficial due to their ubiquity and much larger area of coverage afforded of these devices.

### *7.2.5 Deep Learning for Sleep Estimation*

Zhao et.al have shown in a recent study [124] that deep learning models or (neural network models) can be effectively used for sleep stage estimation using vital signal parameters extracted from electromagnetic signals. They used a combination of convolutional neural network (CNN) model to extract physiological features specific to sleep stages and a recurrent neural network model to extract temporal features pertaining to transition between sleep stages. While the goal of this work was to develop a model that would work across different environmental settings and different participants, in the future, deep learning models can be used more effectively to learn individual specific (within-subject) parameters and can be more accurate over longer periods of time. Researchers have studied within-subject deviations in sleep parameters such as sleep duration, sleep latency and sleep efficiency [74] across 669 participants between the ages of 38-50. They reported that although there were significant variances in these measures nightly, over the course of the year, these variances evened out significantly. Studies observing EEG patterns related to specific sleep stages [25] also reported a clear distinction between high within-subject similarity (i.e. stability), and low between-subject similarity (i.e. variation). These results further call for the need to develop machine learning models that can learn inter-individual traits in sleep and can utilize this information to reflect habitual behavior especially for long-term tracking.

### *7.2.6 Presenting Sleep Feedback*

There has been research done in the HCI community to understand how user perception of data is affected by the visualizations used to represent them [121]. As discussed in Chapter 3, long term sleep tracking devices have the potential to better benefit users if they present them with actionable sleep feedback. However, more research has to be done to figure out how to present this data to users. Findings from the survey on SleepApp indicated that people were more likely to follow recommendations for a positive daytime outcome. This can guide how sleep feedback can be presented to users. Currently, sleep scores and sleep efficiency are presented as a precise, single point value, such as 92%. While these metrics are based on measured used in clinical reports, they are however not interpreted by clinicians but presented directly to users without context.

### *7.2.7 Presenting Objective Metrics in Ranges*

Current sleep sensing technologies infer sleep metrics based on physiological signals such as body movement, breathing, and heart rate. Algorithms that rely purely on physiological bio-markers of sleep cannot as accurately differentiate between awake in bed and asleep in bed as well as an data gathered from an EEG signal. This substitution in sensing modality introduces a certain level of inaccuracy. Therefore reporting sleep scores in ranges rather than single values can give users a better mental model of their sleep data and avoid false precision. Similar to users' perceptions of changes in weight by few pounds [67], daily fluctuations in sleep scores or sleep efficiency does not imply drastic changes in sleep quality and only causes users to be unnecessarily concerned. Systems can provide sleep score ranges instead of a single-point value, based on sleep sensing data and self-assessments. These ranges could focus on overall sleep quality, moving away from representing sleep quality a single value. Doing so will embrace the inherent sensing inaccuracies without compromising on the metrics.

### *7.2.8 Tailoring Feedback on Personal Models*

Another aspect that could be incorporated into sleep quality feedback is supplementing objective sleep metrics with subjective assessment of sleep satisfaction. This can help personalize sleep feedback and help users see how their sleep quality changes compared to their baseline. As discussed in sub-section 7.2.5, inter-individual trait-variability in sleep plays an important role in how much variance in sleep can be seen nightly. Therefore, underlining their sleep metrics when compared to their average or baseline can add more meaning to sleep feedback. This removes the undue emphasis on target sleep scores that maybe population based. Similar approaches to conveying personalized feedback about screen-time or smartphone usage such as the iPhone has been shown to be effective [117]. Screen-time feedback is given weekly in terms of percentage increase or decrease in use compared to the users' average.

### *7.2.9 Enabling Sleep Self Experimentation*

For people suffering from sleep disorders such as insomnia, or poor sleep quality over prolonged periods of time, creating their own n-of-1 self-experiments to identify factors affecting their sleep is crucial to improvement in sleep health. Lauri et. al have created a workbook, No More Sleep Nights [59], to help users create hypothesis based on their answers to a questionnaire and tracking their progress over several weeks. While SleepApp, provides an easy way to track variables affecting sleep and sleep quality, the next step is to provide a self-experimentation platform that allows users to test their hypothesis and confirm their results with data.

### *7.2.10 Reducing User Burden*

In the current SleepApp study, we relied purely on the participant to log their sleep and actively participate in measures to assess alertness and cognitive function. Overtime, this may lead to tracking fatigue as shown in previous work [32]. We can prevent this by incorporating sensor data into the system such as by using a commercial sleep tracker to input objective sleep-related data. Abdullah et al. have proposed in Cognitive Rhythms [17] a novel approach to assess alertness passively from smartphone usage. This can be used as a substitute to the more cumbersome PVT test that participants felt took too long. Another way of reducing user burden can be through finding innovative ways to administer the EMA tests. One such way, proposed by Xioayi et al in [123], could be by incorporating an EMA test, such as answering the subjective sleepiness scale on the lock screen of the smart phone. Feedback from one of the participants in our study was that taking the EMA assessments four times a day was too much. The also felt that if reducing the frequency of the tests from four times a day to three would make it more convenient. Once sufficient data is collected to get a large enough sample size, we can determine if the effect of the outcome is reduced by lowering the frequency of the tests.

### *7.2.11 Measuring Outcomes Such as Stress or Productivity*

There are several daytime carryover effects to sleep in addition to alertness, cognitive function, and mood. Some additional examples include productivity [109], response to stress, or self-control [50]. While these variables are commonly known from personal experiences to be directly impacted by sleep quality, there is potential future work in confirming these from an individual perspective. Exploring ways to incorporate these measures could add another dimension to sleep health.

## *7.3 Lessons Learned*

Through the course of my research work, I have had the unique opportunity to interact with experts across several disciplines. While designing and evaluating the wireless sensors system, I was able to work with engineers as well as technicians in sleep medicine who set-up and monitor polysomnography (PSG) studies as well as those who manually score PSG data. While working on the study for Making Sense of Sleep Sensors, I was able to interact with sleep researchers in the UW nursing school. I became a part of the Quantified Self community on social media and observed interactions in the community around the latest self tracking devices on the market and discussion around informal n-of-1 trials conducted by community members to optimize different aspects of their health. My spouse is an avid self-tracker from whom I got to observe several aspects of the lived informatics model of self tracking. My interactions with with people from these diverse backgrounds has shaped my research goals and well as perspectives.

### *7.3.1 Balancing Need and Novelty*

As engineers, we more often than not gravitate towards novelty and creating new and unique ways to address existing problems. Especially with devices that involve human interaction, novelty can come in the form of altering the form-factor and size of a device, reducing power consumption, enabling wireless capabilities etc. In the amount of time it took me to complete this body of work, several consumer sleep sensing devices have hit the market, risen to popular trends and fallen into bankruptcy or discontinued production. When the entire life-cycle of sleep sensing device manufacturing company seems to be less than a decade, there are some lessons that we need to pay attention to. As is the case with several gadgets, the novelty factor fades away leading to abandonment [79]. As discussed in Chapter 2, the overview of the sleep senors landscape has shown a vast number devices, in different form-factors and sensing modalities. There is a true need for devices that fade into the users environment and are minimally invasive. Especially since these devices are targeted for long-term use, minimizing user burden may lead to longer-periods of engagement.

### *7.3.2 Trade-off Between Comfort and Accuracy*

When it comes to sleep tracking in particular, users prioritize comfort and privacy over accuracy. Several devices with embedded EEG sensors that require users to place a device over their foreheads, have not had been successful despite their higher levels of accuracy. When it comes to long-term sleep tracking for user with no known sleep disorders, offering higher accuracy at the cost of higher intrusiveness may not sit well with users. Device manufacturers need to consider the trade-off between accuracy and comfort and weigh comfort more for this particular use case. Anecdotal evidence as well as supplemental findings from Making Sense of Sleep Sensors has revealed that seemingly insignificant discomforts such as an LED shining from a sleep sensing device worn on the wrist or placed on the bedside can be disruptive to users sleep. While it is well known in the scientific community that sometimes the mere act of observing or sensing a phenomenon can actually alter the phenomenon. This is particularly relevant to sleep. Device manufacturers and designers should pay heed to how their device might potentially alter the very phenomenon it tries to sense: sleep quality.

### *7.3.3 Purpose of Sleep Tracking*

Long-term sleep tracking for users with no sleep disorders can have varied purposes. Sleep is very much a routine based activity for a lot of users and users may not necessarily have much control over it. The timing and duration of sleep can be dictated by the user's occupation, intermittent work deadlines or examinations, life-events such as the addition of a infant or a pet into the family or changes in physical and mental health. The transition from an otherwise satisfactory sleep routine to lowered sleep quality usually bring awareness to users and brings the need for a tool or device to assist in tracking patterns in sleep quality. Herein lies the true power of contextualizing sleep with independent variables and providing meaningful feedback in relation to day-time cognitive function. During these transition periods, users may be more engaged and open to providing EMA data at a higher frequency than when their sleep quality has remained at a sustained level over a long-period of time. Adjusting for these ebbs and flows in user engagement is a vital part of long-term sleep tracking.

### *7.3.4 Users' Intuition Versus Reliance on Technology*

As technologists, we strive to address several fundamentally human problems with technology. Devices and algorithms are interwoven into our lives. Scholars have recently started to uncover negative effects of such high levels of dependence on technology such as the dampening of our own intuition and the way we make sense of our surroundings. This has been noted in avenues like the ability to navigate among Innuit tribes [20] and communication and inter-personal relationships [114]. Sleep quality is also very subjective and personal. Our own sense of satisfaction with sleep and feeling of wellness can carry as much weight as any objective measure. Any technological innovation is best used as an supplement to the human intuition and not a replacement. We must aim to create device and tools that aid users in their pursuit of a healthier and fuller life, and not overpower their intuition with data.

## Chapter 8

# SUMMARY AND CONCLUSION

In this last chapter, I have summarized the contributions in the previous chapters towards my thesis statement.

### *8.1 Summary of Prior Chapters*

In Chapter 1, I discussed the potential of commercial sleep sensing devices for for long term sleep tracking in contrast to validated clinical sleep tracking devices for short term diagnostics. I identified the gaps in current commercial sleep sensing devices, in terms of their hardware and algorithmic limitations, and sleep feedback to facilitate good sleep health . I outlined, through my thesis statement, how understanding the gaps between user perception of sleep data and expert opinion on good sleep practices can inform the design of novel sleep and the provision of meaningful and contextualized feedback can improve sleep health. I described how the studies conducted to answer three main research questions support my thesis statement: (i) How are users impacted by the feedback they receive from current commercially available sleep sensing devices? (ii) How can we enable long-term sleep tracking in users' homes in an unobtrusive manner? (iii) How can we provide actionable sleep feedback that can help users adopt healthy sleep habits?

In Chapter 2, I provided an overview of the different aspects of sleep sensing, sleep feedback and sleep health. I provided a brief description about the physiology of sleep in order to provide context to the sensors that are used to sense the physiological changes that accompany sleep. I then outlined the rigorous sensing protocol that goes into the clinical gold standard of diagnostic sleep sensing: PSG and followed it with progressively unobtrusive sensing solutions provided by commercial devices to enable long-term sleep tracking. I positioned my research contribution in creating a novel sensing solution within the context of the hardware and accuracy limitations of the current commercial devices. I briefly discussed sleep metrics used in the medical community for diagnostic purposes and provided background on user perception of sleep feedback. I surveyed the contributions in the HCI community around sleep tracking, and sleep self-experimentation, to build upon their findings with a specific contribution in providing sleep feedback that is contextualized with day-time cognitive function.

Chapter 3, describes the study conducted to identify gaps between user perception of sleep feedback from commercial devices and expert opinions on best practices for sleep health. The study was designed to answer three research questions (i)how are users making sense of the feedback they receive from commercial sleep sensing devices? (ii) what aspects of sleep sensing and feedback undermine users' ability to achieve good sleep health? and (iii) what aspects of current sleep sensors are in line with evidence-based methods of understanding and promoting sleep health? To answer these questions we collected a data-set consisting of interviews with 5 sleep experts,surveys with 87 and interviews with 12 people that have used sleep sensing devices, and 6986 consumer product reviews from the most widely used commercial sleep sensing devices. Our analysis showed that current sleep sensing devices facilitated good sleep health, especially to those suffering from sleep disorders or chronic conditions that affect sleep, by bringing awareness to users' sleep habits and patterns. Sleep sensors act as barriers due to their inherent sensor limitations leading to inconsistent sleep quality inferences, mismatch between objective data and users' own subjective assessment of their sleep and by providing unactionable sleep feedback. Based on these finding I provided design recommendations for sleep technologists to help provide meaningful and actionable sleep feedback.

In Chapter 4, I outlined the design and implementation of a 2.4GHz wireless system built on Software Defined Radios (SDR): WiBreathe, to physiological signals such as breathing in natural settings in a user's home. This system lays the foundation for a wireless, non-contact, physiological sensor using radio signals in the same range as that of commodity WiFi devices. I explained the theory of operation and the signal processing pipeline to extract the raw breathing waveform from the baseband signal. I further discussed the challenges in implementing the system in the wild and described WiBreathe's adaptive breathing rate estimation ensemble algorithm. I presented specific low SNR (signal to noise ration) scenarios encountered during testing and showed how the performance of WiBreathe's ensemble algorithm is better than any single frequency estimation technique. I described the evaluation of WiBreathe in a controlled setting, showing how the performance is affected by distance and orientation of the user from the TX (transmitter) and RX (receiver) pair. I summarized the results of WiBreathe's evaluation conducted in user's homes, in natural settings. I concluded the chapter with WiBreathe's breathing pattern detection algorithm that can be used to detect events related to breathing disorders especially during sleep.

In Chapter 5, I described the implementation of DoppleSleep: an easily deployable, low-cost, K-Band continuous wave (CW) RADAR for sleep wake detection. This device, being an order of magnitude higher in wavelength than that of the wireless signal implemented in WiBreathe, can detect three physiological bio-markers related to sleep: body movement, breathing rate and heart rate. Although the increase in center frequency affords higher resolution and hence the ability to detect minute chest movements due to hear-beats, it limits the range of detection of the device. However, for sleep-detection, this range is sufficient since the user is confined to their bed during sleep. The base-band signal output from the on-chip mixer is a compound of environmental noise and physiological signals. Since the three physiological signals do not spectrally overlap, multi-stage filtering of the base-band signal gives us the raw time-domain waveform corresponding to these signals. From these signals, we then extract frame level features that are fed into a machine learning model. DoppleSleep, using the Random Forest decision tree algorithm, is able to reliably classify sleep versus wake with 89% precision. The output from the Sleep/Wake classifier was then used to infer objective sleep metrics such as sleep onset latency (SOL), number of awakenings, and total sleep time (TST).

In Chapter 6, I described the design and implementation of SleepApp, a smartphone based application that explores the possibility of providing contextualized and actionable sleep quality feedback by allowing users to track variables that affect their sleep quality and by measuring the day-time cognitive effects of sleep. Through a preliminary study conducted with 9 participants over two weeks, we showed that a liner model can be used to provide users with feedback that shows how their pre-bedtime behaviours would affect specific aspects of their day-time cognitive function such as working memory and alertness. In a follow up survey conducted with over 50 participants, we then explored users' preferences across various behaviours they intended to make changes to and how the decrement or improvement in daytime cognitive function would affect their willingness to make changes to routines. The preliminary findings from this study shows promise for the adoption of contextualized and actionable sleep feedback inorder to improve sleep health.

In Chapter 7, I discussed the limitations of the studies conducted for this thesis work and highlighted opportunities for future work.

### *8.2 Conclusion*

Through my research work, I have provided support for my thesis statement: understanding the gaps between user perception of sleep data and expert opinions on good sleep practices can inform the design of novel sleep sensors and the development of personalized sleep models that can provide users with actionable feedback to improve their sleep health. I have identified and limitations in current commercial sleep sensing devices in the context of intrusiveness for long term sleep tracking and the lack of meaningful feedback to help improve sleep health. I have explored the possibility of using novel wireless sensing techniques for unobtrusive long-term sleep tracking and smartphone application with a suite of journals and EMA tests to provide users with contextualized and actionable feedback. Through an inter-disciplinary approach of addressing multiple aspects of sleep tracking such as sensing, quantifying sleep quality and providing feedback, I have provided a holistic body of work towards more purposeful sleep tracking.

# BIBLIOGRAPHY

[1] Apple watch sleep tracking app. `https://apps.apple.com/app/apple-store/id1138066420`, last accessed on 12/19.

[2] Beddit sleep sensor. `https://www.beddit.com/`, last accessed on 12/19.

[3] Fitbit. `https://www.fitbit.com/home`, last accessed on 12/19.

[4] Global market study on sleep aids: Sleep apnea to witness highest growth by 2020. http://www.persistencemarketresearch.com/marketresearch/sleep-aids-market.asp.

[5] Hexoskin shirt. `https://www.hexoskin.com/`, last accessed on 12/19.

[6] Oura sleep ring. `https://ouraring.com/`, last accessed on 12/19.

[7] Philips smart sleep headband. `https://www.usa.philips.com/c-e/smartsleep.html`, last accessed on 12/19.

[8] Polar sleep plus. `https://www.polar.com/us-en/smart-coaching/polar-sleep-plus`, last accessed on 12/19.

[9] Respironics actiwatch. `http://www.actigraphy.com/solutions/actiwatch/actiwatch2.html`, last accessed on 12/19.

[10] S+ from resmed. `https://sleep.mysplus.com/`, last accessed 12/19.

[11] Samsung galaxy watch sleep tracking. `https://ww.9to5google.com/2018/10/22/galaxy-watch-rem-sleep-tracking-update/`, last accessed on 12/19.

[12] Sense with sleep pill. `https://sleeptrackers.io/sense/`, last accessed 12/19.

[13] Withings aura. `https://support.withings.com/hc/en-us/categories/200189426-Withings-Aura`, last accessed on 12/19.

[14] Withings move. `https://www.withings.com/us/en/withings-move`, last accessed on 12/19.

[15] Zeo head-band. `https://polyphasic.net/polyphasic-sleep/sleep-tracking/zeo-sleep-tracking/`, last accessed on 12/19.

[16] AASM. Sleep Technology:Technical Guideline. Technical report.

[17] Saeed Abdullah, Elizabeth L. Murnane, Mark Matthews, Matthew Kay, Julie A. Kientz, Geri Gay, and Tanzeem Choudhury. Cognitive rhythms. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16*, pages 178–189, New York, New York, USA, 2016. ACM Press.

[18] Torbjorn Åkerstedt, Michel Billiard, Michael Bonnet, Gianluca Ficca, Lucile Garma, Maurizio Mariotti, Piero Salzarulo, and Hartmut Schulz. Awakening from sleep. *Sleep medicine reviews*, 6(4):267–286, 2002.

[19] Sonia Ancoli-Israel, Roger Cole, Cathy Alessi, Mark Chambers, William Moorcroft, and Charles P Pollak. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, 26(3):342–392, 2003.

[20] Claudio Aporta, Eric Higgs, David Hakken, Lisa Palmer, Mark Palmer, Robert Rundstrom, Bryan Pfaffenberger, George Wenzel, Thomas Widlok, Claudio Aporta, et al. Satellite culture: global positioning systems, inuit wayfinding, and the need for a new account of technology. *Current anthropology*, 46(5):729–753, 2005.

[21] Jared Bauer, Sunny Consolvo, Benjamin Greenstein, Jonathan Schooler, Eric Wu, Nathaniel F. Watson, and Julie Kientz. ShutEye. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, page 1401, New York, New York, USA, may 2012. ACM Press.

[22] Adel Belouchrani and Moeness G Amin. Time-frequency music. *IEEE Signal Processing Letters*, 6(5):109–110, 1999.

[23] Sushanth Bhat, Ambra Ferraris, Divya Gupta, Mona Mozafarian, Vincent A DeBari, Neola Gushway-Henry, Satish P Gowda, Peter G Polos, Mitchell Rubinstein, Huzaifa Seidu, et al. Is there a clinical role for smartphone sleep apps? comparison of sleep cycle detection by a smartphone application to polysomnography. *Journal of Clinical Sleep Medicine*, 11(07):709–715, 2015.

[24] Nathália Brandolim Becker, Rute Isabel Santos Martins, Saul de Neves Jesus, Roberto Chiodelli, and Michael Stephen Rieber. Sleep health assessment: A scale validation. *Psychiatry Research*, 259:51–55, jan 2018.

[25] J Buckelmüller, H-P Landolt, HH Stassen, and Peter Achermann. Trait-like individual differences in the human sleep electroencephalogram. *Neuroscience*, 138(1):351–356, 2006.

[26] Daniel J Buysse. Sleep health: can we define it? Does it matter? *Sleep*, 37(1):9–17, jan 2014.

[27] Daniel J. Buysse, Charles F. Reynolds, Timothy H. Monk, Susan R. Berman, and David J. Kupfer. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2):193–213, 1989.

[28] Kun-Mu Chen, Yong Huang, Jianping Zhang, and Adam Norman. Microwave life-detection systems for searching human subjects under earthquake rubble or behind barrier. *IEEE transactions on biomedical engineering*, 47(1):105–114, 2000.

[29] Eun Kyoung Choe, Sunny Consolvo, Nathaniel F. Watson, and Julie A. Kientz. Opportunities for computing technologies to support healthy sleep behaviors. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, page 3053, New York, New York, USA, may 2011. ACM Press.

[30] Eun Kyoung Choe, Bongshin Lee, Matthew Kay, Wanda Pratt, and Julie A. Kientz. SleepTight. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, pages 121–132, New York, New York, USA, sep 2015. ACM Press.

[31] Eun Kyoung Choe, Bongshin Lee, Matthew Kay, Wanda Pratt, and Julie A. Kientz. SleepTight: Low-burden, Self-monitoring Technology for Capturing and Reflecting on Sleep Behaviors. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*, pages 121–132, 2015.

[32] Eun Kyoung Choe, Nicole B Lee, Bongshin Lee, Wanda Pratt, and Julie A Kientz. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1143–1152. ACM, 2014.

[33] David Clifton, J Graham Douglas, Paul S Addison, and James N Watson. Measurement of respiratory rate from the photoplethysmogram in chest clinic patients. *Journal of clinical monitoring and computing*, 21(1):55–61, 2007.

[34] Harvey R Colten, Bruce M Altevogt, , and Institute of Medicine (US) Committee on Sleep Medicine Research. *Sleep Disorders and Sleep Deprivation*. National Academies Press (US), 2006.

[35] Sunny Consolvo, Katherine Everitt, Ian Smith, and James A. Landay. Design requirements for technologies that encourage physical activity. In *Conference on Human Factors in Computing Systems, CHI*, 2006.

[36] Roshan Cools. N-Back Test. pages 822–822, 2010.

[37] Felicia Cordeiro, Daniel A. Epstein, Edison Thomaz, Elizabeth Bales, Arvind K. Jagannathan, Gregory D. Abowd, and James Fogarty. Barriers and Negative Nudges. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, pages 1159–1162, New York, New York, USA, 2015. ACM Press.

[38] John R. Crawford and Julie D. Henry. The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43(3):245–265, sep 2004.

[39] Jesse Dallery, Rachel N Cassidy, and Bethany R Raiff. Single-case experimental designs to evaluate novel technology-based health interventions. *Journal of medical Internet research*, 15(2), 2013.

[40] Nediyana Daskalova, Karthik Desingh, Alexandra Papoutsaki, Diane Schulze, Han Sha, and Jeff Huang. Lessons Learned from Two Cohorts of Personal Informatics Self-Experiments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–22, sep 2017.

[41] Nediyana Daskalova, Bongshin Lee, Jeff Huang, Chester Ni, and Jessica Lundin. Investigating the Effectiveness of Cohort-Based Sleep Recommendations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–19, sep 2018.

[42] Nediyana Daskalova, Danaë Metaxa-Kakavouli, Adrienne Tran, Nicole Nugent, Julie Boergers, John McGeary, and Jeff Huang. SleepCoacher. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16*, pages 347–358, New York, New York, USA, 2016. ACM Press.

[43] Luciane de Souza, Ana Amélia Benedito-Silva, Maria Laura Nogueira Pires, Dalva Poyares, Sergio Tufik, and Helena Maria Calil. Further validation of actigraphy for sleep studies. *Sleep*, 26(1):81–85, 2003.

[44] Massimiliano de Zambotti, Fiona C Baker, Adrian R Willoughby, Job G Godino, David Wing, Kevin Patrick, and Ian M Colrain. Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiology & behavior*, 158:143–149, 2016.

[45] Amy D Droitcour, Olga Boric-Lubecke, Victor M Lubecke, and Jenshan Lin. 0.25/spl mu/m cmos and bicmos single-chip direct-conversion doppler radars for remote sensing of vital signs. In *2002 IEEE International Solid-State Circuits Conference. Digest of Technical Papers (Cat. No. 02CH37315)*, volume 1, pages 348–349. IEEE, 2002.

[46] Amy D Droitcour, Olga Boric-Lubecke, Victor M Lubecke, Jenshan Lin, and Gregory TA Kovacs. Range correlation and i/q performance benefits in single-chip silicon doppler radars for noncontact cardiopulmonary monitoring. *IEEE Transactions on Microwave Theory and Techniques*, 52(3):838–848, 2004.

[47] Daniel A Epstein, Jennifer H Kang, Laura R Pina, James Fogarty, and Sean A Munson. Reconsidering the device in the drawer: lapses as a design opportunity in personal informatics. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 829–840. ACM, 2016.

[48] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, pages 731–742, New York, New York, USA, 2015. ACM Press.

[49] Daniel A Epstein, An Ping, James Fogarty, and Sean A Munson. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 731–742. ACM, 2015.

[50] Liese Exelmans and Jan Van den Bulck. Self-control depletion and sleep duration: the mediating role of television viewing. *Psychology & health*, pages 1–18, 2018.

[51] Guillaume Fuchs, Chamran Ashour, Tom Bäckström, et al. Super-wideband spectral envelope modeling for speech coding. In *Interspeech*, 2019.

[52] Kristina Grifantini. How's my sleep?: Personal sleep trackers are gaining in popularity, but their accuracy is still open to debate. *IEEE pulse*, 5(5):14–18, 2014.

[53] Weixi Gu, Longfei Shangguan, Zheng Yang, and Yunhao Liu. Sleep hunter: Towards fine grained sleep stage tracking with smartphones. *IEEE Transactions on Mobile Computing*, 15(6):1514–1527, 2015.

[54] Weixi Gu, Zheng Yang, Longfei Shangguan, Wei Sun, Kun Jin, and Yunhao Liu. Intelligent sleep stage mining service with smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 649–660. ACM, 2014.

[55] Lauren Hale. Inadequate Sleep Duration as a Public Health and Social Justice Problem: Can We Truly Trade Off Our Daily Activities for More Sleep? *SLEEP*, dec 2014.

[56] Lauren Hale. Inadequate sleep duration as a public health and social justice problem: can we truly trade off our daily activities for more sleep?, 2014.

[57] Tian Hao, Guoliang Xing, and Gang Zhou. isleep: unobtrusive sleep quality monitoring using smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, page 4. ACM, 2013.

[58] Allison G Harvey, Kathleen Stinson, Katriina L Whitaker, Damian Moskovitz, and Harvinder Virk. The subjective meaning of sleep quality: a comparison of individuals with and without insomnia. *Sleep*, 31(3):383–393, 2008.

[59] Peter Hauri. No more sleepless nights workbook. *Tracking Your Progress Toward A Great Night's Sleep*, 2001.

[60] Max Hirshkowitz, Kaitlyn Whiton, Steven M Albert, Cathy Alessi, Oliviero Bruni, Lydia DonCarlos, Nancy Hazen, John Herman, Eliot S Katz, Leila Kheirandish-Gozal, et al. National sleep foundation's sleep time duration recommendations: methodology and results summary. *Sleep health*, 1(1):40–43, 2015.

[61] E. Hoddes, V. Zarcone, H. Smythe, R. Phillips, and W. C. Dement. Quantification of Sleepiness: A New Approach. *Psychophysiology*, 10(4):431–436, jul 1973.

[62] Conrad Iber and Conrad Iber. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*, volume 1. American Academy of Sleep Medicine Westchester, IL, 2007.

[63] Leah A Irish, Christopher E Kline, Heather E Gunn, Daniel J Buysse, and Martica H Hall. The role of sleep hygiene in promoting public health: A review of empirical evidence. *Sleep medicine reviews*, 22:23–36, aug 2015.

[64] Ravi Karkar, Jasmine Zia, Roger Vilardaga, Sonali R Mishra, James Fogarty, Sean A Munson, and Julie A Kientz. A framework for self-experimentation in personalized health. *Journal of the American Medical Informatics Association*, 23(3):440–448, 2015.

[65] Matthew Kay, Eun Kyoung Choe, Jesse Shepherd, Benjamin Greenstein, Nathaniel Watson, Sunny Consolvo, and Julie A. Kientz. Lullaby. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, page 226, 2012.

[66] Matthew Kay, Tara Kola, Jessica R Hullman, and Sean A Munson. When (ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems. In *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems (CHI '16)*, 2016.

[67] Matthew Kay, Dan Morris, Julie A Kientz, et al. There's no such thing as gaining a pound: Reconsidering the bathroom scale user interface. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 401–410. ACM, 2013.

[68] Matthew Kay, Dan Morris, Mc Schraefel, and Julie a. Kientz. There's No Such Thing as Gaining a Pound: Reconsidering the Bathroom Scale User Interface. *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13*, pages 401–410, 2013.

[69] Matthew Kay, Kyle Rector, Sunny Consolvo, Ben Greenstein, Jacob Wobbrock, Nathaniel Watson, and Julie Kientz. PVT-Touch: Adapting a Reaction Time Test for Touchscreen Devices. In *Proceedings of the ICTs for improving Patients Rehabilitation Research Techniques*. IEEE, 2013.

[70] Matthew Kay, Kyle Rector, Sunny Consolvo, Ben Greenstein, Jacob Wobbrock, Nathaniel Watson, and Julie Kientz. PVT-Touch: Adapting a Reaction Time Test for Touchscreen Devices. In *Proceedings of the ICTs for improving Patients Rehabilitation Research Techniques*, pages 248–251. IEEE, 2013.

[71] Alan E Kazdin. Reactive self-monitoring: the effects of response desirability, goal setting, and feedback. *Journal of consulting and clinical psychology*, 42(5):704, 1974.

[72] Sharon Keenan and Max Hirshkowitz. Monitoring and Staging Human Sleep. *Principles and Practice of Sleep Medicine: Fifth Edition*, pages 1602–1609, 2010.

[73] Hannah K. Knudsen, Lori J. Ducharme, and Paul M. Roman. Job stress and poor sleep quality: Data from an American sample of full-time workers. *Social Science & Medicine*, 64(10):1997–2007, may 2007.

[74] Kristen L Knutson, Paul J Rathouz, Lijing L Yan, Kiang Liu, and Diane S Lauderdale. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the cardia study. *Sleep*, 30(6):793–796, 2007.

[75] Ping-Ru T Ko, Julie A Kientz, Eun Kyoung Choe, Matthew Kay, Carol A Landis, and Nathaniel F Watson. Consumer sleep technologies: a review of the landscape. *Journal of clinical sleep medicine*, 11(12):1455–1461, 2015.

[76] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM computer communication review*, volume 45, pages 269–282. ACM, 2015.

[77] Meir H. Kryger, T. (Tom) Roth, and William C. Dement. *Principles and practice of sleep medicine*. Saunders/Elsevier, 2011.

[78] Steven Lanzisera, David T Lin, and Kristofer SJ Pister. Rf time of flight ranging for wireless sensor network localization. In *2006 international workshop on intelligent solutions in embedded systems*, pages 1–12. IEEE, 2006.

[79] Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H Nguyen. Why we use and abandon smart devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 635–646. ACM, 2015.

[80] Jisoo Lee, Erin Walker, Winslow Burleson, Matthew Kay, Matthew Buman, and Eric B. Hekler. Self-Experimentation for Behavior Change. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pages 6837–6849, New York, New York, USA, 2017. ACM Press.

[81] Changzhi Li and Jenshan Lin. Complex signal demodulation and random body movement cancellation techniques for non-contact vital sign detection. In *2008 IEEE MTT-S International Microwave Symposium Digest*, pages 567–570. IEEE, 2008.

[82] Changzhi Li, Victor M Lubecke, Olga Boric-Lubecke, and Jenshan Lin. A review on recent advances in doppler radar sensors for noncontact healthcare monitoring. *IEEE Transactions on microwave theory and techniques*, 61(5):2046–2060, 2013.

[83] Changzhi Li, Yanming Xiao, and Jenshan Lin. Experiment and spectral analysis of a low-power *ka*-band heartbeat detector measuring from four sides of a human body. *IEEE Transactions on Microwave Theory and Techniques*, 54(12):4464–4471, 2006.

[84] Changzhi Li, Yanming Xiao, and Jenshan Lin. Design guidelines for radio frequency non-contact vital sign detection. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1651–1654. IEEE, 2007.

[85] Ian Li, Anind Dey, and Jodi Forlizzi. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 557–566. ACM, 2010.

[86] Ian Li, Anind K Dey, and Jodi Forlizzi. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 405–414. ACM, 2011.

[87] Qin Li and Les Atlas. Time-variant least squares harmonic modeling. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03).*, volume 2, pages II–41. IEEE, 2003.

[88] Zilu Liang, Bernd Ploderer, and Mario Alberto Chapa-Martell. Is fitbit fit for sleep-tracking?: sources of measurement errors and proposed countermeasures. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 476–479. ACM, 2017.

[89] Life Sciences & Transformational Health. Global Market Study on Sleep Aids: Sleep Apnea to Witness Highest Growth by 2020. Technical report.

[90] James C Lin. Noninvasive microwave measurement of respiration. *Proceedings of the IEEE*, 63(10):1530–1530, 1975.

[91] Wanyu Liu, Bernd Ploderer, and Thuong Hoang. In bed with technology: challenges and opportunities for sleep tracking. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, pages 142–151. ACM, 2015.

[92] Bram Lohman, Olga Boric-Lubecke, VM Lubecke, PW Ong, and MM Sondhi. A digital signal processor for doppler radar sensing of vital signs. *IEEE Engineering in Medicine and Biology Magazine*, 21(5):161–164, 2002.

[93] F Charles Mace and Thomas R Kratochwill. Theories of reactivity in self-monitoring: A comparison of cognitive-behavioral and operant models. *Behavior Modification*, 9(3):323–343, 1985.

[94] Thomas Melia, Scott Rickard, and Conor Fearon. Histogram-based blind source separation of more sources than sensors using a duet-esprit technique. In *2005 13th European Signal Processing Conference*, pages 1–4. IEEE, 2005.

[95] JOCHEN MEYER, JUDY KAY, DANIEL A EPSTEIN, PARISA ESLAMBOLCHILAR, and LIE MING TANG. A life of data-characteristics and challenges of very long-term self-tracking for health and wellness. 2019.

[96] Jun-Ki Min, Afsaneh Doryab, Jason Wiese, Shahriyar Amini, John Zimmerman, and Jason I. Hong. Toss 'n' turn. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pages 477–486, New York, New York, USA, apr 2014. ACM Press.

[97] Hawley E. Montgomery-Downs, Salvatore P. Insana, and Jonathan A. Bond. Movement toward a novel activity monitoring device. *Sleep and Breathing*, 16(3):913–917, sep 2012.

[98] Inbal Nahum-Shani, Eric B Hekler, and Donna Spruijt-Metz. Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*, 34(S):1209, 2015.

[99] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Nathaniel Watson. Contactless sleep apnea detection on smartphones. In *Proceedings of the 13th annual international conference on mobile systems, applications, and services*, pages 45–57. ACM, 2015.

[100] J Adam Noah, Carol Boliek, Tania Lam, and Jaynie F Yang. Breathing frequency changes at the onset of stepping in human infants. *Journal of neurophysiology*, 99(3):1224–1234, 2008.

[101] Adrian A. Ong and M. Boyd Gillespie. Overview of smartphone applications for sleep analysis. *World Journal of Otorhinolaryngology-Head and Neck Surgery*, 2(1):45–49, mar 2016.

[102] Byung-Kwon Park, Olga Boric-Lubecke, and Victor M Lubecke. Arctangent demodulation with dc offset compensation in quadrature doppler radar receiver systems. *IEEE transactions on Microwave theory and techniques*, 55(5):1073–1079, 2007.

[103] A C Parrott and I Hindmarch. The Leeds Sleep Evaluation Questionnaire in psychopharmacological investigations - a review. *Psychopharmacology*, 71(2):173–9, 1980.

[104] John P. Pollak, Phil Adams, and Geri Gay. PAM. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, page 725, New York, New York, USA, 2011. ACM Press.

[105] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. Whole-home gesture recognition using wireless signals. In *Proceedings of the 19th annual international conference on Mobile computing & networking*, pages 27–38. ACM, 2013.

[106] Ruth Ravichandran, Elliot Saba, Ke-Yu Chen, Mayank Goel, Sidhant Gupta, and Shwetak N Patel. Wibreathe: Estimating respiration rate using wireless signals in natural settings in the home. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 131–139. IEEE, 2015.

[107] Peng Rong and Mihail L Sichitiu. Angle of arrival localization for wireless sensor networks. In *2006 3rd annual IEEE communications society on sensor and ad hoc communications and networks*, volume 1, pages 374–382. Ieee, 2006.

[108] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. Personal tracking as lived informatics. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1163–1172. ACM, 2014.

[109] Mark R Rosekind, Kevin B Gregory, Melissa M Mallis, Summer L Brandt, Brian Seal, and Debra Lerner. The cost of poor sleep: workplace productivity loss and associated costs. *Journal of Occupational and Environmental Medicine*, 52(1):91–98, 2010.

[110] Thomas Roth, Catherine Coulouvrat, Goeran Hajak, Matthew D Lakoma, Nancy A Sampson, Victoria Shahly, Alicia C Shillington, Judith J Stephenson, James K Walsh, and Ronald C Kessler. Prevalence and perceived health associated with insomnia based on dsm-iv-tr; international statistical classification of diseases and related health problems, tenth revision; and research diagnostic criteria/international classification of sleep disorders, criteria: results from the america insomnia survey. *Biological psychiatry*, 69(6):592–600, 2011.

[111] Avi Sadeh and Christine Acebo. The role of actigraphy in sleep medicine. *Sleep Medicine Reviews*, 6(2):113–124, 2002.

[112] Anita Valanju Shelgikar, Patricia F Anderson, and Marc R Stephens. Sleep tracking, wearable technology, and opportunities for research and clinical care. *Chest*, 150(3):732–743, 2016.

[113] Masayuki Shimohira, Toshihide Shiiki, Junko Sugimoto, Yukiko Ohsawa, Michio Fukumizu, Takeshi Hasegawa, Yoshihide Iwakawa, Yoshiko Nomura, and Masaya Segawa. Video analysis of gross body movements during sleep. *Psychiatry and clinical neurosciences*, 52(2):176–177, 1998.

[114] Catherine Steiner-Adair and Teresa H Barker. *The big disconnect: Protecting childhood and family relationships in the digital age*. Harper Business, 2013.

[115] Edward J Stepanski and James K Wyatt. Use of sleep hygiene in the treatment of insomnia. *Sleep medicine reviews*, 7(3):215–25, jun 2003.

[116] Giulio Tononi and Chiara Cirelli. Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, 10(1):49–62, feb 2006.

[117] Ebubechukwu Onyekachi Aham Ubochi. *Apple's Screen Time: The Utility of Digital Wellness Software in Curbing Smartphone Dependency*. PhD thesis, 2019.

[118] P. A. Van Dongen, Maurice D. Baynard, Greg Maislin, and David F. Dinges. Systematic Interindividual Differences in Neurobehavioral Impairment from Sleep Loss: Evidence of Trait-Like Differential Vulnerability. *Sleep*, 27(3):423–433, may 2004.

[119] Joris C. (Joris Cornelis) Verster, S. R. Pandi-Perumal, and David L. Streiner. *No Title*. Humana Press, Totowa, NJ, 2008.

[120] Yanming Xiao, Jenshan Lin, Olga Boric-Lubecke, and M Lubecke. Frequency-tuning technique for remote detection of heartbeat and respiration using low-power double-sideband transmission in the ka-band. *IEEE Transactions on Microwave Theory and Techniques*, 54(5):2023–2032, 2006.

[121] Huahai Yang, Yunyao Li, and Michelle X Zhou. Understand users' comprehension and preferences for composing information visualizations. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(1):6, 2014.

[122] Rayoung Yang, Eunice Shin, Mark W Newman, and Mark S Ackerman. When fitness trackers don't'fit': end-user difficulties in the assessment of personal tracking device accuracy. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 623–634. ACM, 2015.

[123] Xiaoyi Zhang, Laura R Pina, and James Fogarty. Examining unlock journaling with diaries and reminders for in situ self-report in health and wellness. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5658–5664. ACM, 2016.

[124] Mingmin Zhao, Shichao Yue, Dina Katabi, Tommi S Jaakkola, and Matt T Bianchi. Learning sleep stages from radio signals: A conditional adversarial architecture. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 4100–4109. JMLR. org, 2017.