The Traumschreiber System

Enabling Crowd-based, Automated,

Complex, Polysomnographic

Sleep and Dream Experiments

Dissertation

Kristoffer Appel

Abstract

Sleep and dreaming are important research topics. Unfortunately, the methods for researching them have several shortcomings. In-laboratory polysomnographic sleep and dream research is a costly, time-consuming and effortful endeavor, often resulting in small subject counts. Moreover, the unfamiliar sleeping environment can lead to distorted measurements as compared to the natural sleep environment at the subject's home.

Conducting sleep and dream experiments in the field by a crowd of subjects could be a solution. However, complex experiment paradigms cannot be investigated this way, because there are no tools available, which enable naive subjects to carry out complex polysomnographic studies on their own.

The *Traumschreiber* system, which is developed and evaluated in this dissertation, offers a solution to this problem. It consists of a high-tech sleep mask and a minicomputer, and enables naive crowd subjects to perform complex polysomnographic sleep and dream experiments at home. On the one hand, it instructs the crowd subject, what to do when. On the other hand, it controls the experiment during the time the subject is asleep, analyzing the data in real-time using state-of-the art machine learning techniques. The rationale behind is to enable a big data approach to sleep and dream research, using the data recorded by a crowd of subjects for large-scale investigations about sleep and dreaming, with low costs for the researcher.

After describing the development process of the Traumschreiber system, its usefulness regarding crowd-based automated polysomnographic field studies is evaluated. First, it is validated against a commercial medical polysomnographic sleep laboratory system, demonstrating its good polysomnographic data recording capabilities – including measurements of EEG, EOG, EMG and ECG –, which enable the researcher to identify typical sleep patterns like slow waves or rapid eye movements as well as sleep stages in the recorded data.

Furthermore, two field studies show, that the Traumschreiber system can be used successfully by naive subjects to conduct complex sleep experiments at their homes. This includes acoustic stimulation of the sleeping subject as well as sleep stage dependent activities of the system. The sleep staging algorithm implements a Keras/Tensorflow based neural network approach, which demonstrates the system's readiness for state-of-the-art machine learning techniques. However, the currently used neural network is kept very simple and can determine the sleep stage not very reliably; it should be further developed and trained on more data of more subjects.

The Traumschreiber system will be made available under an open source license, enabling any researcher to use, modify or further develop it. A description, how to produce arbitrarily many entities of the Traumschreiber system, is given in this dissertation and shows that one system can be produced at low costs in a short amount of time (production costs per piece around 140 EUR plus 80 EUR for the other parts of the system, production time: one hour per item).

Taken together, a new tool for sleep and dream research is now available, which enables a crowd-based approach of gathering polysomnographic data from complex sleep and dream experiments with only low resource needs for the scientist.

Table of Contents

Abstract....3

Figures....7

Tables....9

Abbreviations....10

Acknowledgements....11

Preface – Reflecting on my scientific activities....13

1 Introduction....17
- 1.1 The basic ideas....17
- 1.2 The importance and variety of sleep and dream research....19
- 1.3 Assessment of sleep and dreaming in the sleep laboratory....20
- 1.4 Conducting sleep experiments as field studies at the subjects' homes....24
  - 1.4.1 (Online) questionnaire studies....24
  - 1.4.2 Field studies using electronic tools....27
    - 1.4.2.1 Portable medical sleep recording devices....27
    - 1.4.2.2 Sleep recording consumer products....29
- 1.5 Artificial intelligence....32
  - 1.5.1 Brief historical overview....33
  - 1.5.2 Machine learning and deep neural networks....35
- 1.6 Sleep and dream research and artificial intelligence....39
- 1.7 Aim of this dissertation....40

2 Methods....41
- 2.1 Overview....41
- 2.2 Development....41
  - 2.2.1 Defining the requirements towards the Traumschreiber system....41
  - 2.2.2 Iteration-based development....43
- 2.3 Evaluation....50
  - 2.3.1 First study: Sleep laboratory validation....50
    - 2.3.1.1 Aim....50
    - 2.3.1.2 Subjects....50
    - 2.3.1.3 Materials and procedure....50
    - 2.3.1.4 Data analysis....51
  - 2.3.2 Second study: Home-based prototype study....52

2.3.2.1 Aim....52
2.3.2.2 Subjects....52
2.3.2.3 Materials....53
2.3.2.4 Procedure....54
2.3.2.5 Arousal threshold experiment....55
2.3.2.6 Data analysis....55
2.3.3 Third study: Preparing polysomnographic crowd experiments....56
2.3.3.1 Aim....56
2.3.3.2 Description of the preparation steps for a Traumschreiber system study....57
2.3.4 Fourth study: A crowd-based polysomnographic experiment....58
2.3.4.1 Aim....58
2.3.4.2 Subjects....59
2.3.4.3 Materials....59
2.3.4.4 Time frame....59
2.3.4.5 Procedure....60
2.3.4.6 Data analysis....62

3 Results....66
3.1 Development....66
3.1.1 Overview....66
3.1.2 The components of the Traumschreiber system....69
3.1.2.1 Hardware....69
3.1.2.1.1 Sleep mask....69
3.1.2.1.2 Experiment station....71
3.1.2.2 Software....73
3.1.2.2.1 Autonomous, easy to program experiments....73
3.1.2.2.2 Automatic sleep staging and pattern recognition....75
3.1.2.2.3 Real-time data processing using encapsulated, modular scripts....77
3.1.2.2.4 Data visualization and data analyses....77
3.2 Evaluation....88
3.2.1 First study: Sleep laboratory validation....88
3.2.1.1 Direct comparison of both systems....89
3.2.1.2 Sleep staging comparison....90
3.2.2 Second study: Home-based prototype study....91
3.2.2.1 Data quality....91
3.2.2.2 Effect of the system on sleep and dreaming....92
3.2.2.3 Ease of use and comfort of the sleep mask....94
3.2.2.4 Effect of acoustic stimulation on the occurrence of arousal....94
3.2.2.5 How to further improve the system....95

3.2.3 Third study: Preparing polysomnographic crowd experiments....96
3.2.3.1 Description of production work step and cost analysis....96
3.2.3.2 Programming the experiment....107
3.2.4 Fourth study: A crowd-based polysomnographic experiment....108
3.2.4.1 General results....108
3.2.4.2 Data quality....109
3.2.4.3 Sleep quality....115
3.2.4.4 Ease of use, comfort and durability of the Traumschreiber system....115
3.2.4.5 Feedback of the subjects to the sleep mask and the experiment....116
3.2.4.6 Automatic sleep scoring algorithm and subject-specific adjustment....116
3.2.4.7 N3 and REM sleep differences....118
3.2.4.8 Sleep depth....119
3.2.4.9 Workload of the experimenter and costs of the experiment....120

4 Discussion....121
4.1 General discussion....121
4.2 Discussion of the development process and evaluation studies....123
4.2.1 Discussion of the development process....123
4.2.2 Evaluation studies....124
4.2.2.1 First study: Sleep laboratory validation....126
4.2.2.2 Second study: Home-based prototype study....127
4.2.2.3 Third study: Preparing polysomnographic crowd experiments....128
4.2.2.4 Fourth study: A crowd-based polysomnographic experiment....129
4.3 Future research directions and prospects....132
4.3.1 Future research directions....132
4.3.1.1 Further developing the Traumschreiber system....132
4.3.1.2 Suggestions for further studies with the Traumschreiber system....135
4.3.2 Future prospects....137

5 References....139

Appendices....150

Figures

Figure 1.1: Screenshot of an online study navigation...26
Figure 1.2: An exemplary sleep experiment online questionnaire......................................26
Figure 1.3: Overview of open source fields...32
Figure 1.4: Example of an autonomous car controlled by an artificial intelligence.............34
Figure 1.5: Image labels generated by an artificial neural network....................................37
Figure 1.6: Features learned by so called convolutional neural networks..........................38
Figure 1.7: Example of a picture created by an artificial neural network............................38
Figure 2.1: Requirements towards the Traumschreiber sleep mask...................................42
Figure 2.2: Requirements towards the Traumschreiber experiment station........................42
Figure 2.3: Requirements towards the Traumschreiber data analysis................................42
Figure 2.4: Picture of the technical development process...44
Figure 2.5: Electric prototype version 0.2..47
Figure 2.6: Picture of the textile development seminar...47
Figure 2.7: One intermediate result of the textile development process.............................47
Figure 2.8: The first self-made prototype of the textile part of the sleep mask....................47
Figure 2.9: Traumschreiber sleep mask version 1.0...48
Figure 2.10: Traumschreiber sleep mask 1.0 together with the eye cover..........................48
Figure 2.11: Traumschreiber sleep mask version 2.2...48
Figure 2.12: The final version of the Traumschreiber sleep mask (open pocket)................49
Figure 2.13: The final version of the Traumschreiber sleep mask (closed pocket).............49
Figure 2.14: Sticky electrodes of two systems on the forehead..51
Figure 2.15: Screenshot of the instruction video (1)..54
Figure 2.16: Screenshot of the instruction video (2)..61
Figure 2.17: Screenshot of the instruction video (3)..61
Figure 2.18: The experiment xml file of the fourth study..62
Figure 3.1: The final version of the Traumschreiber sleep mask...66
Figure 3.2: One complete Traumschreiber system experiment box....................................67
Figure 3.3: N3 sleep recorded autonomously by the Traumschreiber system.....................68
Figure 3.4: N2 sleep recorded autonomously by the Traumschreiber system.....................68
Figure 3.5: REM sleep recorded autonomously by the Traumschreiber system..................68
Figure 3.6: Time frequency plots of a whole night recording...69
Figure 3.7: Four screenshots of the prototype Android app..72
Figure 3.8: Example xml experiment file..73
Figure 3.9: Example of the real-time pattern detection..76
Figure 3.10: Signal transmission plot (good transmission quality)......................................79
Figure 3.11: Signal transmission rate plot (bad transmission quality).................................80
Figure 3.12: The GUI of the post_visualization.py script...81

Figure 3.13: Night overview plot..82
Figure 3.14: Detailed view plot..82
Figure 3.15: The events dropdown menu...83
Figure 3.16: Specifying a filter in the config xml file...84
Figure 3.17: Manual sleep scoring with the Traumschreiber system..................................85
Figure 3.18: Real-time sleep stage predictions of the Traumschreiber system..................86
Figure 3.19: The same predictions, averaged using a 20 minute sliding window...............87
Figure 3.20: Polysomnographic data recorded with the ALICE system..............................88
Figure 3.21: The same data recorded with the Traumschreiber system.............................89
Figure 3.22: Sleep staging comparison between the two systems....................................91
Figure 3.23: Sleep quality in the first night and the control night.......................................92
Figure 3.24: Feeling of recovery after the experimental night and the control night...........93
Figure 3.25: Comparison of acoustic stimuli regarding arousals..95
Figure 3.26: Back and front side of the equipped PCB..97
Figure 3.27: All parts of 20 complete Traumschreiber systems..100
Figure 3.28: The battery holder before cutting...101
Figure 3.29: The battery holder after cutting..101
Figure 3.30: Screenshot of the video tutorial (1)..103
Figure 3.31: Sewing pattern for the sleep mask bag...104
Figure 3.32: Screenshot of the video tutorial (2)..105
Figure 3.33: Example of the bio signal calibration procedure...109
Figure 3.34: Example of a nightly arousal procedure (1)..110
Figure 3.35: Example of a nightly arousal procedure (2)..110
Figure 3.36: Example of REM sleep recorded by the Traumschreiber system.................112
Figure 3.37: Unfiltered, raw data example of a recording with bad quality........................112
Figure 3.38: Example of a recording with good data quality...113
Figure 3.39: Example of a recording with bad data quality...113
Figure 3.40: Recording success during the second and third night..................................114
Figure 3.41: Automatic sleep staging comparison...117
Figure 3.42: Difference between REM and N3 sleep dreams (1).....................................118
Figure 3.43: Difference between REM and N3 sleep dreams (2).....................................118
Figure 3.44: Difference between REM and N3 sleep dreams (3).....................................119
Figure 3.45: Difference between REM and N3 sleep dreams (4).....................................119
Figure 3.46: Difference between REM and N3 sleep..119

Tables

Table 1.1: Comparison of consumer sleep recording devices..31
Table 2.1: Overview of the main electric development iterations.......................................46
Table 3.1: Overview of the electrode positions and recorded channels..............................70
Table 3.2: Detailed description of the experiment xml file commands.................................74
Table 3.3: Polysomnographic graphoelements in both systems..89
Table 3.4: Reasons were for having trouble falling asleep...94
Table 3.5: Overview of the work steps for producing a Traumschreiber system.................96
Table 3.6: Recordings with a good EEG, EMG, EOG, and ECG signal............................114

Abbreviations

AI	Artificial Intelligence
ADC	Analog-to-Digital Converter
BLE	Bluetooth Low Energy
CSV	Comma-Separated Values
ECG	Electrocardiography
ECTS	European Credit Transfer System
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
EUR	Euro
FoR	Feeling of Recovery
FFT	Fast Fourier Transform
GUI	Graphical User Interface
HEOG	Horizontal Electrooculography
IC	Integrated Circuit
ICSD-3	International Classification of Sleep Disorders – Third Edition
InAmp	Instrumentation Amplifier
JSON	JavaScript Object Notation
LED	Light-Emitting Diode
N1	Sleep Stage N1
N2	Sleep Stage N2
N3	Sleep Stage N3
NDR	Norddeutscher Rundfunk
OpAmp	Operational Amplifier
PCB	Printed Circuit Board
Ph.D.	Doctor of Philosophy
POSIX	Portable Operating System Interface
RAM	Random-access Memory
REM	Rapid Eye Movement Sleep
SF-A	Schlaffragebogen A (Sleep Questionnaire A)
SQ	Sleep Quality
SWS	Slow Wave Sleep
UART	Universal Asynchronous Receiver Transmitter
UDP	User Datagram Protocol
USB	Universal Serial Bus
USD	US Dollar
VEOG	Vertical Electrooculography
XML	Extensible Markup Language

Acknowledgements

I would like to thank quite a number of people, who supported me during my Ph.D. time.

First of all, my supervisor, Prof. Dr. Pipa, from whom I learned a lot about doing science in general and about neuroinformatics, about advanced EEG analysis, about the “hidden” tasks a researcher has to do like networking and about budgeting. Moreover, I’d like to thank him for financing my Ph.D. position and parts of my experiments, even though sleep and dream science is not his main research interest, and for all the freedom throughout my Ph.D. and “letting me do my thing”.

Then, the other members of the PhD board, for their willingness to read and evaluate this dissertation: Prof. Dr. Kühnberger, Prof. Dr. Schredl and Prof. Dr. Dresler.

Next, I would like to thank all my co-workers in and around the office and the sleep laboratory, for valuable discussions and the good working atmosphere, and especially Mr. Leugering, who played a major role in developing the electro-technical parts of the high-tech sleep mask discussed in this dissertation and parts of the software to steer it.

Then, all the students, who worked together with me in one of the several projects, on which this dissertation is built. First, the 10 bachelor students writing their thesis in the field of sleep and dreaming, with me as their supervisor. Second, all the students from the Student Initiative Sleep & Dream, who not only conducted a complete sleep laboratory study with 20 subjects in their leisure time, but also helped building up a very active sleep & dream student community in Osnabrueck, with excursions to far away countries, movie nights, lucid dreaming workshops and more. Third, the master students of the study project on Sleep and Memory, which I supervised, for all the efforts they put into conducting the first sleep laboratory study of this kind in Osnabrueck. Fourth, the students of my seminars “Sleep & Dream”, “Reading Club: Dreams and Dreaming”, and “Textile Design of a Hightech Sleep Mask I + II”, who also contributed to the success of this dissertation with inspiration and ideas, and Prof. Dr. Hähnel for the joint supervision of the textile design courses.

Next, I would like to thank the scientists from my personal research network, for sharing their knowledge with me, the inspiring discussions, and the great time at conferences, workshops and other meetings.

I would also like to thank Prof. Dr. Steiger and his group at the Max Planck Institute of Psychiatry in Munich for the opportunity to learn, how sleep research is conducted in a professional sleep laboratory, during my stay there as a visiting researcher.

Moreover, I would like to thank Prof. Dr. Young from the Muenster University Hospital, for his support, the opportunity to collaborate in a joint study, and especially the mediation of the donation of a professional sleep laboratory equipment to the Osnabrueck Student Initiative Sleep & Dream. Also I would like to thank Prof. Dr. König for letting me and the other students use his sleep research equipment.

Furthermore, I would like to thank the members of the Osnabrueck University ethics committee for their helpful advice and consulting during in total four ethics proposals.

Next, I would like to thank Mrs. Jungeilges and Mrs. Eibisch for their help with all the administrative things, which need to be done in German bureaucracy.

Also, I would like to thank the Hans-Mühlenhoff-Stiftung zur Förderung leistungsfreudiger Jugend and the German Sleep Society for their financial support of my experiments, in total worth nearly 10.000 EUR, without which some studies of this dissertation could not have been conducted.

And, last, but not least, I would like to thank all my friends and family for the continuous support throughout the last years.

Thank you!

Preface – Reflecting on my scientific activities

The highest academic degree, the Ph.D., is awarded for scientific achievements and proves that the holder of the degree is able to conduct elaborated scientific works on his or her own[1]. Before starting with the actual dissertation, I would like to briefly reflect on my scientific activities during the past four years.

It all started by following a personal interest of mine – lucid dreaming. I started researching it in my leisure time using a low-budget one-channel sleep EEG device. One year later, I created a student initiative about sleep and dreaming at my university in 2012. I motivated fellow students to conduct our own research in the field of sleep and dreaming using portable high-tech sleep devices, applied for funding from the general students' committee, and organized regular meetings. Subsequently, I wrote my master's thesis about how it is possible to communicate with (sleeping!) lucid dreamers in real-time.

Fortunately, Prof. Dr. Pipa offered me the possibility to extend my research during a Ph.D.. I organized two university rooms and raised a large equipment donation of a medical device manufacturer, enabling me to conduct research in my own fully-equipped sleep laboratory. I went to the Max Planck Institute of Psychiatry in Munich as a visiting researcher, where I took part in a research project on lucid dreaming using a high-density 128 channel EEG, and learned, how to conduct professional sleep and dream research. This knowledge I then transferred to my sleep laboratory in Osnabrueck, and passed it on to fellow students of the student initiative, as well as during seminars I taught. Besides conducting research, life in academia also includes teaching. I offered courses worth 40 ECTS points in total; sometimes more than 100 students participated in my seminars. I supervised 19 bachelor's and master's theses and a one-year study project of ten master students.

The before mentioned student initiative grew steadily and became the largest young scientists sleep and dream research group in Germany, if not worldwide. We conducted workshops and experiments in the sleep laboratory and went on three scientific excursions, e. g. one week in spring 2016 to sleep researchers in Finland. More details

1 Osnabrueck University Ph.D. regulations "Promotionsordnung des Fachbereichs Humanwissenschaften der Universität Osnabrueck für die Verleihung des Grades Ph.D. in Cognitive Science" (2011), § 1.

about the student initiative can be found in the article “Kurzvorstellung der Studentischen Initiative Sleep & Dream Osnabrueck”, published in the journal “Somnologie” (Appel, 2017, see appendix).

There were several research projects of mine, which I conducted during the last years. I validated the sleep communication from my master’s thesis using polysomnographic recordings in my sleep laboratory at Osnabrueck university with experiment subjects from all over Germany. I recorded and analyzed lucid dreams with a high-density EEG at the Max Planck Institute Institute of Psychiatry in Munich. I conducted an online study about the success rate of a lucid dream induction technique, for which I used a self-developed website as a platform. I analyzed a dietary supplement regarding its effect on sleep and dreaming. I carried out a sleep laboratory study investigating the role, which sleep plays for memory processes. I conducted another questionnaire study about the public perception of lucid dreams. I successfully replicated the results of a recent study about a new lucid dreaming induction technique in my sleep laboratory, and started a study investigating the effects of lucid dreaming smartphone apps. I wrote a review paper about lucid dreaming, and an article about the same topic for the “DGSM Kompendium Schlafmedizin” of the German Sleep Society, which is used for educating purposes of sleep laboratory staff and medical doctors. I summarized a German dissertation about event-related potentials during lucid dreams and published it in a scientific journal in order to make it available to the English speaking scientific community. In many of these projects, I instructed and supervised undergraduate students, who carried out the experiments together with me as a part of their theses. Other smaller projects are not mentioned here.

I made most of this research available to the scientific community, e. g. at scientific conferences. Some of the research is already published in scientific journals as well, with the other not yet published works to follow after finishing my Ph.D.. A list of publications and conference presentations is attached at the end of this preface. Besides publishing my research in the scientific literature, it was important to me to transfer the new knowledge to the non-scientific public as well. I presented my research twice on television (NDR), in an international pop-sci magazine (The New Scientist), was interviewed by a large German internet magazine (Netzpiloten), and by journalists from Finland to Brazil (see appendix for examples).

Finally, I carried out other scientific tasks throughout my Ph.D.. From obtaining funds for my experiments (in total nearly 10.000 EUR), to writing and defending four complete ethics proposals, to acting as a reviewer in peer-reviewed journals four times. I connected the new Osnabrueck sleep and dream research to well-known researchers from various institutes, such as the Donders Institute (Nijmegen, The Netherlands), the Central Institute of Mental Health (Mannheim), the Max Planck Institute of Psychiatry (Munich), the Charité Sleep Competence Center (Berlin), the Muenster University Hospital, and Bern University (Switzerland), and cooperated with them in common research projects.

One project was not yet named in this reflection of my scientific activities: the Traumschreiber project. It is the main project of my Ph.D., and is described in larger detail in this dissertation.

List of scientific publications and conference presentations

Appel, K., Pipa, G., & Dresler, M. (2017). Investigating consciousness in the sleep laboratory - an interdisciplinary perspective on lucid dreaming. *Interdisciplinary Science Reviews,* 1-16.

Appel, K., Pipa, G. (2017). Auditory evoked potentials in lucid dreams: a dissertation summary. *International Journal of Dream Research*, 10(1), 98-100.

Appel, K. (2017). Kurzvorstellung der Studentischen Initiative Sleep & Dream Osnabrueck. *Somnologie*, 21(3), 252-254.

Kern, S., Appel, K., Schredl, M., & Pipa, G. (2017). No effect of alpha-GPC on lucid dream induction or dream content. *Somnologie*, 21(3), 180-186.

Appel, K., Leugering, J., Pipa, G. (2016). Kommunikation mit Klarträumern in Echtzeit – aktuelle Möglichkeiten. Oral presentation at the 24rd Annual Conference of the German Sleep Society, Dresden, Germany.

Leugering, J., Appel, K. (2016). Traumschreiber: Eine Open Source Schlafmaske, die Hirnaktivitäten messen und beeinflussen kann. Oral presentation at the MetaRheinMainChaosDays of the Chaos Computer Club, Darmstadt, Germany.

Appel, K., Leugering, J., Pipa, G. (2016). “Traumschreiber”: Measuring and manipulating human sleep with a portable high-quality but low-cost polysomnographic system. Poster presentation at the 23rd Congress of the European Sleep Research Society, Bologna, Italy.

Appel, K. (2016). The Student Initiative Sleep & Dream Osnabrueck. Oral presentation at the 33rd Annual International Dream Conference, Kerkrade, Netherlands.

Appel, K., Kern, S. (2015). Phänomenologie luzider Träume und Induktionstechniken. Schulz, Geisler, Rodenbeck (Hrsg.) *DGSM Kompendium Schlafmedizin.* Ecomed Verlag, Landsberg, XVIII-3.9.1

Appel, K. (2015). The Student Initiative Sleep & Dream Osnabrueck. Oral presentation at the 23rd Annual Conference of the German Sleep Society, Mainz, Germany.

Appel, K., Pipa, G. (2015). Sleep Communication. Poster presentation at the Worldsleep Conference, Istanbul, Turkey.

Kern, S., Appel, K., Schredl, M., Pipa, G. (2015). Lucid dream induction using L-alpha glycerylphosphorylcholine. Poster presentation at the Worldsleep Conference, Istanbul, Turkey.

Appel, K., Pipa, G. (2014). Bidirectional, voluntary and conscious communication with arbitrary message content between a sleeping person and the waking world is possible. Poster presentation at the 22rd Annual Conference of the German Sleep Society, Cologne, Germany.

Appel, K., Erlacher, D., Pipa, G. (2014). Sleep Communication. Poster presentation at the Interdisciplinary College, Guenne, Germany.

1 Introduction

1.1 The basic ideas

All humans sleep on a more or less regular basis; sleep is one of the basic human needs. Moreover, everybody dreams, and even though some people have more trouble remembering their nightly hallucinatory experiences than others, dreaming is a core element of everybody's sleep (Nir & Tononi, 2010).

If sleep or dreaming are disordered, this leads to undesired effects to the individual well-being on the one hand. On the other hand, the society as a whole is affected negatively by sleep disorders: The economic damage to the German national economy due to sleep disorders, for example, accumulated to 1.027 billion EUR in the year 2015 (Federal Statistical Office Germany (Destatis), 2017).

Since everybody sleeps and dreams, everybody is a potential subject for scientific sleep and dream studies. Thus, a crowd-based approach to researching this topic seems to be an idea worth investigating – not only since the more "classical" sleep and dream research methods (e. g. in-lab recordings) suffer from several limitations, as will be shown later in this introduction. The term "crowd-based" as used in this dissertation describes the idea to use the efforts and data of masses of subjects. The aim is to "crowd-source" (as in "outsource") many of the tasks conducted by professional sleep researchers to the naive subjects, so that by far more data can be collected using this parallelized approach than would be possible the standard way. Crowd-based sleep and dream research would not necessarily mean, that the crowd would have to consist of many semi-professional sleep experts. It would be enough to enable a large group of naive subjects to conduct complex, state-of-the-art sleep and dream experiments with professional, polysomnographic methods on their own at home with easy-to-use tools.

Why was this not done until now? Actually, several approaches to measure sleep in larger subject populations outside the laboratory have been undertaken in the past, as will be described in greater detail later in this chapter – ranging from online questionnaire studies to dream diary studies to actigraphy studies. Without going too much into detail here already, it can be summarized that these approaches have several advantages and several disadvantages over measuring sleep in a sleep laboratory. Unfortunately, a crowd-based approach is obviously not well suited for sleep laboratory studies, as conducting in-laboratory research is an endeavor with high resource needs. At the same time, complex sleep experiments, which manipulate the sleep acoustically, which rely on sleep stage-dependent paradigms or which need other ways of interaction between the subject and the experimenter, cannot be conducted by naive subjects on their own at home – until now.

Recent technological developments could be a game changer. Everybody knows, that electric components are getting smaller, cheaper and at the same time more powerful. The knowledge, how to construct complex technical devices, is made available in open source projects on the internet, e. g. for building an electroencephalograph (EEG) in the OpenEEG project[2]. More and more powerful algorithms are developed in the field of artificial intelligence and one of its core disciplines, machine learning, as will be described later in this introduction.

Combining these developments, a system can now be built, which guides naive crowd subjects through a sleep or dream experiment and which conducts parts of it on its own. Such a small, transportable system would have to be capable of explaining to the inexperienced crowd subject, which experiment step to do when, for example how to record the brain activity using electroencephalography (EEG). Moreover, and equally important, it would have to be capable of conducting some parts of the experiment autonomously on its own, e. g. during the time the subject is asleep. Lastly, in an ideal case, setting up scientific experiments and programming this system would be very simple and not require a lot of time or programming knowledge of the sleep scientist. As a result, large-scale, crowd-based sleep and dream research becomes possible.

In the following theoretical part of this dissertation, the backgrounds of modern sleep and dream research will be characterized: its variety, its methodology and limitations. Afterwards, the technological progress in the field of artificial intelligence and machine learning will be briefly recapped, and it will be shown, how these technologies are used in the scientific study of sleep and dreaming already today. Ultimately, building on top of this theoretical work, a system will be developed and evaluated, which makes possible crowd-based polysomnographic sleep and dream experiments: the Traumschreiber system.

2 www.openeeg.sourceforge.net

1.2 The importance and variety of sleep and dream research

Scientific research about human sleep and dreaming is an important field with many facets, as can be seen from the following examples:

- In the field of memory consolidation, scientific sleep studies explore, how memory processes are influenced by sleep, e. g. how sleep helps to better memorize vocabulary, to consolidate freshly learned motor tasks, or to even rearrange existing memories and, thus, gain new insights into current problems. Moreover, it is researched, how these memory processes can be influenced and supported, for example using acoustic stimulation during sleep (for an overview, see Diekelmann (2014)).
- In the field of chronobiology, sleep researchers investigate, which biological processes cause sleep and wakefulness, how sleep is influenced by light exposure, by shift work, or by trans-meridian travel, how jet-lag can be reduced, and other related topics (Sack et al., 2007).
- Recurrent nightmares are just one out of around 60 sleep disorders, which the International Classification of Sleep Disorders lists (ICSD-3, American Academy of Sleep Medicine, 2014). About 5% of the population suffer from recurrent nightmares (Schredl, 2013), and a multiple of this suffer from sleep disorders in general, such as insomnia, sleep related breathing disorders, narcolepsy, or restless legs syndrom (Luyster, Strollo, Zee, & Walsh, 2012), underlining the importance of experimental human sleep research for clinical and therapeutic purposes.
- Lucid dreaming research is one of the most spectacular and fascinating subfields of sleep science – especially to laymen –, as can be seen by the occurrence of this topic in popular scientific journals and television shows (see examples for this in the appendix). Lucid dreams, in which the dreamer realizes that he or she is dreaming whilst staying asleep, offer fantastic opportunities not only for hedonistic purposes, but also as a tool used in several other academic disciplines, ranging from philosophy to sports science to psychotherapy (Appel, Pipa, & Dresler, 2017).

Depending on the exact matter of investigation and the available resources, sleep can be assessed in multiple different ways.

1.3 Assessment of sleep and dreaming in the sleep laboratory

One way to assess sleep – and connected to it, dreaming – is to record it in a scientific sleep laboratory. The standard method is to conduct polysomnographic (from Greek “polus”: many, much, Latin “somnus”: sleep, Greek “graphein”: to write) measurements of a subject’s sleep, i. e. to record multiple physiological sleep parameters: brain activity measured by electroencephalography (EEG), muscle tension at the chin measured by electromyography (EMG), and eye movements measured by electrooculography (EOG).

Recording these parameters allows for specialized data analysis, for example, a classification of discrete 30-second time intervals into the commonly used sleep stages, which helps to get an overview of the macrostructure of the sleep of a specific subject or larger populations: Besides the stage Wake, there is, on the one hand, rapid eye movement (REM) sleep, on the other hand, the non-REM sleep phases N1 (which mainly describes the transition period directly after falling asleep), N2 (“light” sleep), and N3 (“deep” sleep, also called slow-wave-sleep, SWS) (Rama & Zachariah, 2013). Standardized rules have been developed for classification purposes (Rechtschaffen & Kales, 1968; Iber, Ancoli-Israel, Chesson, & Quan, 2007), which are usually applied to the recorded data by manual (visual) scoring, or, lately, by using automatic sleep staging algorithms (Penzel et al., 2007). However, the historically evolved approach of using 30 second based sleep stage bins is criticized, as it has many drawbacks and as the digitalization of sleep recordings makes more advanced analyses possible, for example frequency-based analyses of the data using wavelets or Fourier transformations (Schulz, 2008).

When determining the above mentioned macrostructure of sleep, the occurrence of local patterns in the polysomnographic recordings is analyzed, so called graphoelements, which form the microstructure of the sleep. To these belong sleep spindles, K-complexes, slow waves, rapid eye movements, and several more (Rama & Zachariah, 2013).

Furthermore, often additional polysomnographic characteristics are recorded, such as the heart rate or the heart rate variability using electrocardiography (ECG), movements of the legs, respiration, oxygen saturation levels, infrared camera videos, body movements, or snoring (Rama & Zachariah, 2013)).

Studies about the effect of auditory, visual or other stimulation of sleep on the sleep itself, on the occurrence of arousals or on cognitive performance during following waking phases belong to the first scientific experiments of modern sleep research (Davis, Davis, Loomis, Hervey, & Hobart, 1939; Dement & Wolpert, 1958) and are still a widely used tool in modern sleep and dream research (for a literature review, see (Braumann, 2016)[3]).

The recorded data are then used to answer specific research questions in order to find out more about sleep or dreaming, both from theoretical as well as practical viewpoints, e. g. "Do sleep spindles and occur more often after learning a declarative task?" (Gais, Mölle, Helms, & Born, 2002), or, "How does the intake of alcohol affect the sleep macrostructure and the sleep fragmentation?" (Williams, MacLean, & Cairns, 1983).

Polysomnographic sleep laboratory studies have a number of disadvantages. One of them is the limited availability of research-oriented polysomnographic sleep laboratories with adequate technical and personnel equipment (Ghegan, Angelos, Stonebraker, & Gillespie, 2006), not only in Germany and other Western countries, but to a much larger extent in less developed countries.

Furthermore, conducting polysomnographic sleep laboratory experiments is an effortful, resource-demanding and time-consuming activity (Van De Water, Holmes, & Hurley, 2011), as is illustrated in the following excerpt of the paper "Investigating consciousness in the sleep laboratory - an interdisciplinary perspective on lucid dreaming" (Appel, Pipa, & Dresler, 2017, full text in the appendix):

> *"3. A night in the sleep laboratory*
>
> *[…] Subjects with frequent lucid dreams, which in addition have to be long enough to conduct scientific experiments within them, are hard to find. It is thus not uncommon to invite participants from the whole country or even from abroad for a lucid dreaming study, requiring effort for advertisement, recruitment, payment and traveling.*
>
> *Preparation and performance of actual lucid dreaming experiments in the sleep laboratory are complex as well - besides the usual extensive procedures necessary to conduct sleep laboratory studies, including as a minimum*

3 Bachelor's thesis of Sophia Braumann, supervised by the author of this dissertation. See online appendix.

polysomnographic measurements of the brain activity by EEG, of eye movements by electrooculography (EOG), and of muscle activity by electromyography (EMG). For these measurements, electrodes are fixed on the scalp and in the face of the subject. […] As a result, preparation for the experiment can take several hours before the subject is sent to bed. In lucid dreaming studies with high-density EEG for topographical analyses or source localization of brain activity, for example, it takes two experimenters about three hours to instruct and prepare the subject, including fitting of the 128 EEG electrodes and measuring their positions and orientations.

Then the actual sleep recordings begin: In contrast to daytime EEG studies, where the subject usually sits still in front of a computer screen, sleeping subjects are turning around frequently, moving their heads, rubbing them into the pillow, pulling cables, and so on. Thus, lucid dreaming experimenters usually need to stay awake for the whole night in order to continuously supervise the quality of the recordings and to intervene in case of bad signal quality. There is nothing more frustrating for a lucid dream researcher than a bad EEG signal during the critical few minutes of a lucid dream. […]

The necessity to wake up subjects in case of a lucid dream without self-initiated awakening requires the experimenter to continuously monitor the eye recordings on the computer monitor throughout the whole night. While unfortunately still no automatized solutions for lucid dreaming LRLR eye signal detection in the sleep laboratory exist, some researchers have recently started developing such tools by applying methods from the field of machine learning and pattern recognition to automatically detect lucidity-indicating eye movements or bad EEG signals in real-time (e. g. Appel, 2013).

In the morning, when the subject declares that he or she cannot sleep any more, recordings are stopped, and it usually takes another one or two hours to let the subject fill out more questionnaires, to tidy up and clean the equipment, or to take a few other measures depending on the experiment. For example, a classical study protocol for motor skill training during lucid dreaming needs performance measurements of the chosen skill both in the evening and in the morning. All in all, the whole experimental procedure of recording one night *including preparations and follow-up usually takes from 8 PM until 11 AM the next day – for a few recording minutes of data of interest in the best case scenario, i. e. the subject experiencing an actual lucid dream. In many cases, two or three consecutive nights are recorded in a row in order to reduce traveling demands on the subject, resulting in an extensive workload for the lucid dream researcher. Given all these efforts and constraints, it is not surprising that lucid dreaming studies often suffer from a limited number of participants – if exceeding the character of case-studies of single subjects at all."*

Even though this kind of sleep laboratory study is probably one of the extreme ones – in most cases subjects can be invited from the local area and often most of the recorded nights result in useful data – this insight into a lucid dreaming sleep laboratory study illustrates the immense resource needs of sleep laboratory studies:

- Time-wise for the experimenter or a commissioned person (e. g. a night watch), as he or she has to a) spend the evening in the laboratory preparing the experiment, b) stay awake the whole night in order to supervise the signal quality and to conduct the experimental tasks at night, such as collecting dream reports, applying stimuli, or waking up the subject at specific time points, c) in the morning for uncabling the subject, for collecting more questionnaires, and for cleaning up.
- Time-wise for the subject, as it a) has to spend the evening, night, and (sometimes) morning in the sleep laboratory, instead of sleeping at home, b) needs to travel to the experiment location and back, which might be in a different city, c) might be sleep deprived and thus needs more sleep the following night, due to the unfamiliar clinical sleep laboratory surrounding.
- Money-wise, as the sleep laboratory needs to be heated, cleaned, needs electricity and water, for washing the blankets, replacing materials from time to time, and – especially – for all the sleep laboratory equipment, which can cost several thousand EUR. Moreover, the personnel has to be paid, i. e. the night watch, cleaning personnel, and eventually additional research assistants.
- Infrastructure-wise, as suited rooms have to be allocated for sleep research use.

Besides all these high resource investments, it might, even worse, be the case, that the sleep laboratory environment influences the sleep to such an extent, that it is not comparable to the “natural” sleep at the subject’s home anymore, which one originally planned to study. One reason for this is, that individual needs of the subject such as their preferred time to go to bed are often hard to meet in the standardized procedures of sleep laboratory recordings. Another reason is the so called “first night effect”, which states, that the first night of sleep in a sleep laboratory usually contains more awake periods, and less REM sleep, than in consecutive nights (Agnew, Webb, & Williams, 1966). Hobson, Pace-Schott, & Stickgold (2000) explain nicely, why sleeping in the laboratory is less comfortable, more difficult and less deep for the subject:

> *“To appreciate this point, the reader need only imagine going to an unfamiliar place in an inner city neighborhood of dubious safety, encountering a technician who is a stranger and often of the opposite sex, having ten electrodes affixed to the scalp with cement that smells like airplane dope and then being bid ‘goodnight’ and ‘pleasant dreams.’”*

Furthermore, as Waterman, Elton, & Kenemans (1993) argue, can differences between home and in the laboratory recorded dreams occur. This is due to environmental factors such as the artificial setting or due to the interpersonal, communicative relationship between the experimenter and the subject in the laboratory, which might influence qualitative and quantitative aspects of dream recall and report.

Hobson et al. (2000) suggest, that there are three ways to deal with these limitations of sleep laboratory studies: a) by extending the number of nights for each subject to around seven, which has the disadvantage of increasing the efforts severely, b) by recording only one to four nights, and thus accept possible influences on sleep architecture and dreams, or c) to conduct the experiments as field studies at the subjects’ homes, using portable technology.

1.4 Conducting sleep experiments as field studies at the subjects’ homes

1.4.1 (Online) questionnaire studies

The simplest form of gaining data about sleep is the use of questionnaires or sleep diaries. These can either be used on their own, e. g. for collecting data in epidemiological studies, or in combination with further instructions for the subject, for example when assessing the effect of behavioral tasks on sleep.

As Goril & Shapiro (2013) point out, the main reasons for using standardized questionnaires and rating scales are a) that they are quick and accurate for finding a population of interest, e. g. with a specific sleep disorder such as recurrent nightmares, b) that they enable the researcher to monitor change and progress over longer periods of time, c) that they are quick and easy to evaluate and, in the case of standardized questionnaires, deliver validated results regarding the target they have been developed for, and d) that they allow for direct comparison of subject groups between different study locations and laboratories, for example in multi-site experiments. Moreover, they are relatively inexpensive in comparison to sleep laboratory studies, at least if no license fees have to be paid, and need far less resources.

There exist more than one hundred standardized sleep scales for assessing various kinds of sleep habits and disorders (Shahid, Wilkinson, Marcu, & Shapiro, 2012). Three exemplary and widely used scales in sleep research are

- the *Pittsburgh Sleep Quality Index* (PSQI), which measures the sleep quality and disturbances over a period of one month (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989),
- the *Epworth Sleepiness Scale* (ESS), which measures daytime sleepiness (Johns, 1991),
- the *Morningness-Eveningness Questionnaire* (MEQ), which assesses the chronotype of a subject (Horne & Ostberg, 1976).

A German language questionnaire, which is designed for a differential diagnosis of sleep disorders and is used by the Osnabrueck sleep research group for excluding sleep disordered subjects from study participation, is the *Landecker Inventar zur Erfassung von Schlafstörungen* (LISST). The LISST is based on the *International Classification of Sleep Disorders* (ICSD) and discriminates between sleep disorders such as sleep-related breathing disorders, insomnia, narcolepsy, restless legs syndrom, and circadian rhythm disorders (Schürmann et al., 2001).

Besides using standardized questionnaires, it is of course also possible to develop own questionnaires based on the research question one wants to investigate, however, with the disadvantage that it is not possible to compare the results to norms then.

Questionnaire based studies can be conducted in an online fashion over the internet, saving time and resources. One example of such an internet based study with self-developed questionnaires (conducted by the author of this dissertation) is depicted in figures 1.1 and 1.2.

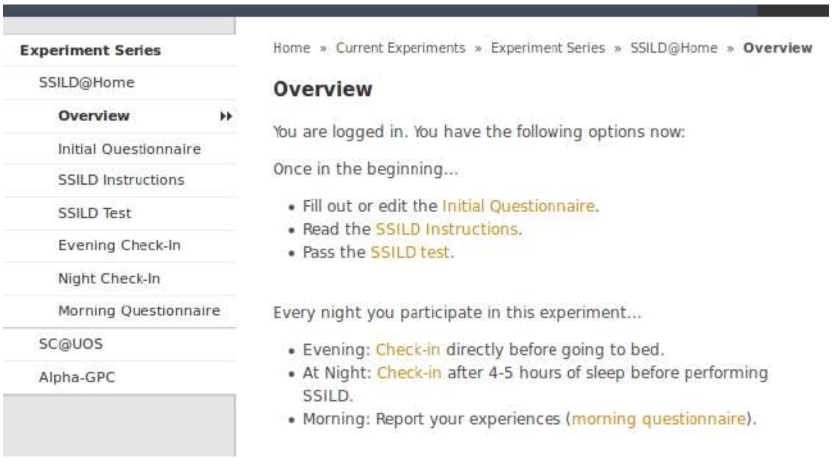


Figure 1.1: Screenshot of an online study navigation, instructing the subject what to do when.

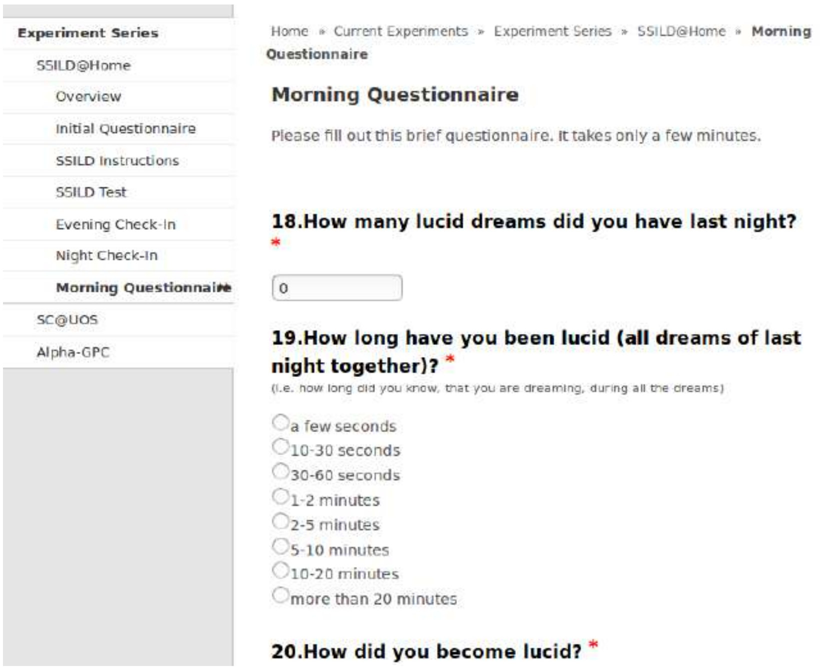


Figure 1.2: An exemplary sleep experiment online questionnaire.

It is possible to extend the idea of questionnaire-based home studies, by supplying the subjects with further experimental materials, such as substances for intake (e. g. drugs or dietary supplements), measurement tools (e. g. saliva sampling) or stimulating tools (e. g.

glasses filtering out blue light to be worn in the evening). These tools can be of electronic nature or non-electronic. Electronic tools are discussed in the next section of this chapter.

One exemplary questionnaire based field study analyzed the effect of a special type of self-administered dietary supplements on sleep and dreaming (Kern, Appel, Schredl, & Pipa, 2017, see appendix)[4]. In this experiment, the subjects had to wake themselves up using an alarm clock after a given time period, open a letter containing either placebo or verum pills, continue sleeping, and, in the morning, report their dreams. The whole procedure was then repeated two more times. Even though this study is rather complex, it has to be noted, that drawing conclusions about sleep stages or other sleep characteristics – except the subjectively experienced ones assessed by the questionnaires – was not possible with such a design.

1.4.2 Field studies using electronic tools

Electronic devices exist, which measure sleep outside the laboratory environment at the subject’s home. There is a large variation, ranging from consumer products (smartphone apps, watches, but also devices capable of recording sleep EEG), to medical devices such as actigraphs and even portable polysomnographic tools similar to those in the sleep laboratory.

1.4.2.1 Portable medical sleep recording devices

Actigraphy and portable polysomnographic devices are the two main technologies in the field of medical sleep recording devices for home use, even though other approaches like bed sensors are developed as well (Van De Water et al., 2011).

Actigraphy is one of the most widely used medical devices in home-based sleep research (Slater et al., 2015). In most cases, actigraphy consists of a watch, which the subject wears around the wrist day and night, even though other places on the subject's body have been tried out with less recording success, too (Zinkhan et al., 2014). The watch contains an accelerometer, which measures arm movements in all directions. Most modern actigraphy devices measure additional parameters as well, for example brightness or pulse. It is possible, to infer from the recorded data, whether the subject is asleep or awake – at least to some degree (Marino et al., 2013). Actigraphy is often used to study

4 Originally conducted as the bachelor’s thesis of Simon Kern, supervised by the author of this dissertation.

circadian rhythms or sleep illnesses like the shift work disorder. However, it is not recommended as a diagnostic tool for routine diagnoses (Kushida et al., 2005; Morgenthaler et al., 2007). The advantages of actigraphy as compared to in-laboratory polysomnography are, that it is less expensive, less complicated to use, ubiquitously available, and most likely without a first night effect (Ancoli-Israel et al., 2003). The disadvantages of actigraphy are, that it cannot measure the brain activity (EEG), eye movements (EOG) or muscle tension (EMG), that it tends to overestimate the amount of sleep and underestimates the amount of wakefulness, and that not all devices have been validated against polysomnography (Ancoli-Israel et al., 2003; Lichstein et al., 2006; Sadeh & Acebo, 2002).

Portable polysomnographic devices, which measure EEG, EOG, EMG and often also ECG, are available, too, and can be used successfully for unattended sleep recordings at the subject's home (Bruyneel & Ninane, 2014). Subjects usually cannot put on the devices on their own (Myllymaa et al., 2016), a trained sleep technician is necessary for connecting the subject to the device in the evening, which is a major disadvantage in comparison to actigraphy or to the consumer devices described below. Moreover, this technology is more expensive, complex and time-consuming than other portable sleep recording technologies. The main advantage over other portable sleep recording devices is that it allows complete sleep evaluation including EEG, EOG, EMG and often more. Compared to in-laboratory polysomnography, the subjects' sleep tends to be better at home, and the costs are lower, however, the data quality is slightly worse (Campbell & Neill, 2011).

Only few data are available regarding the material costs and the cost of actually using home-based polysomnographic devices compared to in-laboratory recordings (Bruyneel & Ninane, 2014). Based on the expenses at the University Osnabrueck sleep laboratory and based on personal communication with several exhibitors at the sleep technology exhibition of the 24th Annual Meeting of the German Sleep Society, 2016, it can be estimated, that state-of-the-art actigraphy watches can easily cost more than 600 EUR, a complete polysomnographic equipment costs several thousand to tens of thousand EUR, with portable polysomnographic sleep recording devices lying in between. The costs for personnel and other expenses are not included in this, but can reach several hundred EUR per night as well. However, this strongly depends on the country, in which the recording takes place (Bruyneel & Ninane, 2014).

The before described medical sleep recording devices are sold commercially, are proprietary (closed) technology and do not allow real-time access of the raw data, displaying further disadvantages of this type of technology.

1.4.2.2 Sleep recording consumer products

A large number of consumer products, which claim to measure human sleep characteristics, are available on the market or preparing a market entry. Even though it was not able to actually measure sleep, the NovaDreamer sleep mask (based on the DreamLight sleep mask described in (LaBerge, Levitan, Rich, & Dement, 1988)) was probably the first commercially available technical sleep consumer product. Its purpose was to influence the subject's sleep and to induce lucid dreams using visual stimulation (blinking lights).

It took until 2009, when the first consumer product was developed, which could actually measure sleep. Furthermore, it was capable of sleep staging the data in real-time and was even sleep laboratory validated: the Zeo system (Griessenberger, Heib, Kunz, Hoedlmoser, & Schabus, 2013; Shambroom, Fábregas, & Johnstone, 2012). The Zeo system consisted of a wireless headband with textile electrodes, which recorded a one-channel prefrontal EEG, and a small station, which received the data in nearly real-time and calculated and displayed the sleep stages. Moreover, it was possible to access the raw data by connecting a cable to the station, and it was very inexpensive (around 150 EUR). However, due to a not working business model, the Zeo went out of business in 2013 (Orlin, 2013).

Nowadays, several products have filled the gap, which was left after the Zeo shutdown. Leaving smartphone apps and fitness trackers aside, which are very popular, but at the same time have very poor quality for measuring sleep and wakefulness (Kolla, Mansukhani, & Mansukhani, 2016), these can be divided into two branches. On the one hand, there are devices, which mainly aim at actual sleep recording. On the other hand, consumer products exist, which mainly aim at measuring brain activity in general, but could be used for measuring sleep as well. Both types of devices cost around 300-800 USD[5] and most of them are produced by small startups, which are financed via crowdfunding campaigns.

---

5 Please see the companies' websites for the latest price.

The consumer products, which are targeted at measuring brain activity during wakefulness or mediation, are not suited well for recording sleep data: Either because they are uncomfortable to wear a whole night long due to a rigid corpus (OpenBCI[6], Emotiv EPOC+[7]), because they are not designed for unattended use during sleep experiments in the field (OpenEEG[8]) or because the batteries last only around three hours (Muse[9]). A sleep recording subproject exists for the OpenBCI, which aims at enabling professional sleep research experiments with the device. However, the costs for setting up this sleep measuring version are comparatively high (almost 950 USD)[10] and actually measuring sleep data is complicated, if not impossible for not inexperienced subjects.

Ten consumer devices could be identified after an extensive internet research, which claim to be able to record sleep data (see table 1.1). Most of them focus on lucid dream induction, relaxation in general, or improving the sleep quality. However, as of November 2017, most of them are not (yet) available. It has to be noted, that the development of most of the devices, which are not available until the date of writing this, was delayed by months or even years. It can thus only be speculated, when these devices can be bought. Furthermore, it has to be noted, that crowdfunding campaigns were canceled in at least two cases before the development of the devices was finished, even though several hundred thousand EUR were collected.

Major disadvantages of the consumer sleep recording devices are, that the data quality is questionable, especially since most of them record only one or two electrophysiological channels. Moreover, nearly all of the devices for conducting sleep experiment measurements at the subject's home are sold by commercial companies for profit[11]. As a result, the systems are proprietary technology, i. e. not open, which means that the exact way, how they function, is kept secret. Consequently, the researcher has to believe the manufacturer about what is being measured and cannot know in detail, how the signal is processed (e. g. filtered) inside the system. Additionally, a way to access to the raw data in

6 www.openbci.com

7 www.emotiv.com

8 www.openeeg.sourceforge.net

9 www.choosemuse.com

10 www.spisop.org/openbci

11 This is especially the case for the medical portable devices, and also applies to sleep laboratory systems.

real-time is available only in a few cases, and also the software for analyzing the data is often proprietary. This makes it often impossible for researchers to adapt the software to specific experimental needs, which exceed the standard data processing pipeline.

Name **Website**	**Available?**	**Price (including discounts)**	**Recording channels**	**Sleep stimulation built in?**	**Raw data accessible in real-time?**	**Software open source?**	**Sleep laboratory validated?**	**Comment**
LdreamM *ldreamm.com*	yes	360 USD	8x EEG	magnetic stimulation, auditory stimulation	yes	no	no	Seems to be a one-man project of a lucid dreaming and meditation hobbyist, with a small community of users. Uses wet EEG electrodes, which are complicated to use by naive subjects. The project has an unscientific touch, e. g. the developer describes measuring energy levels during meeting his dead wife's soul during wakefulness.
Dreem *dreem.com*	no	399 USD	6x EEG, accelerometer, pulse oximeter	auditory	no	no	no	Extensive testing with hundreds of subjects took place. Not financed via crowdfunding. Strong research collaborations ongoing. Unclear shipment date (postponed several times).
Aurora *sleepwithaurora.com*	no	299 USD	at least 1x EEG, claims to measure EOG, EMG and ECG as well, accelerometer	visual	No (not specified)	no	no	Unclear shipment date, delayed for years since 2014.
iBand+ *ibandplus.com*	no	309 USD	1x EEG	visual, auditory	yes	yes	no	A special developer version has to be bought for accessing raw data in real-time. Unclear shipment date (postponed several times).
kokoon *kokoon.io*	no	239 EUR	1x EEG, accelerometer, microphone	auditory	no	no	no	Unclear shipment date (postponed several times).
neuroonOpen *neuroonopen.com*	no	199 USD	1x EEG, pulse oximeter, thermometer, accelerometer	visual, auditory	yes	yes	no	There was a previous version before (not produced anymore). Unclear shipment date of the new version.
LucidCatcher *kck.st/2qoxywq*	no	380 USD	2x EEG	electric stimulation	No (not specified)	No (not specified)	no	Shipment date unclear (development since 2014). Developer version costs extra.
Lucid Dreamer *luciddreamer.com*	no	-	-	-	-	-	-	canceled due to ineffectiveness regarding lucid dream induction
LUCI *(no website)*	no	-	-	-	-	-	-	canceled, fraud allegations mentioned
iBrain *neurovigil.com*	no	-	1x EEG	-	-	-	-	Device is not commercially available for everybody, but only to special partners. Several prominent people support this device.

Table 1.1: Comparison of consumer sleep recording and stimulation devices.

Even though no exceptions to this are available regarding devices designed for actual sleep recording, the OpenBCI project and the OpenEEG project show an alternative way, by open sourcing their hardware specifications and software code. Open hardware and open software in this context mean, that all the knowledge needed for reproducing the system is made publicly available, even though specific licenses can apply e. g. for commercial usage. Similar open developments take place in other contexts as well (see figure 1.3), for example in the scientific literature (open access journal articles) or in projects like Wikipedia (open knowledge).

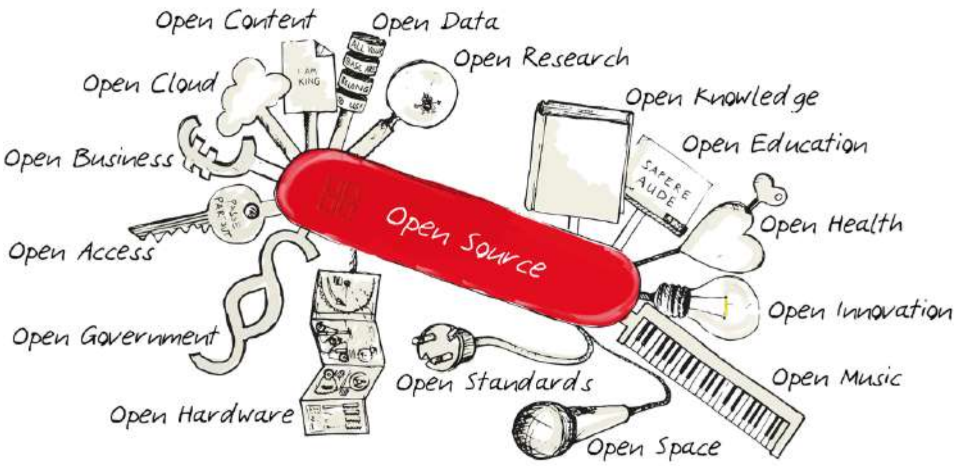


Figure 1.3: Overview of open source fields. Author: Johannes Spielhagen, Bamberg, Germany - CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=27179850

1.5 Artificial intelligence

If a researcher wants to let naive crowd subjects conduct complex, interactive polysomnographic sleep or dream experiment at their home, he or she cannot just hand out sleep recording devices to the subjects. Instead, each subject needs to be instructed how to use the device, and someone has to control the experiments during the time the subject is asleep. For example, in a sleep stage dependent experimental paradigm, in which the subject is woken up during a specific sleep stage in order to fill out a questionnaire, someone needs to classify the sleep data in real-time.

Artificial intelligence (AI) is a buzzword, which is used in all kinds of contexts. As the Stanford *One Hundred Year Study on Artificial Intelligence* states, there are many definitions of what AI is, with none being universally accepted (Stanford University, 2016).

The authors write that “a useful one” by (Nils J. Nilsson, 2009) is the following: “Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment.”

Transfered to machines, which are used in complex polysomnographic sleep and dream experiments, AI could be used to take over those tasks, which are normally conducted by a human intelligence (in this case, the sleep researcher or night watch). For example, an AI, which features machine learning algorithms, could be trained to detect sleep stages and to carry out specified experiment tasks based on this, replacing the human intelligence by an artificial counterpart. Before describing, in how far AI is used in the sleep and dream research context nowadays already, some background information on AI and one of its most important subfields, machine learning, is given.

1.5.1 Brief historical overview

The foundations of AI have been laid a long time ago (for a more detailed overview, see for example Russell, S. & Norvig (1995)). Philosophers like Aristotle, who lived in the ancient Greece, already discussed the idea, that the mind is in some ways like a machine, which operates on knowledge encoded in an internal language. By using thought, it can be decided, which actions should be taken. Later, mathematicians like George Boole, Thomas Bayes or Gottfried Leibniz started to develop tools to manipulate statements of logical certainty as well as uncertain, probabilistic statements, and the idea to use algorithms for reasoning purposes was born. Again several centuries later, scientists from other academic fields such as psychology and linguistics contributed to the foundations of AI by adding insight about the mechanisms of human intelligence, already demonstrating this area of research’s multidisciplinarity.

The term *artificial intelligence* itself was first coined in the proposal for the famous Dartmouth workshop written in the year 1955 (McCarthy, Minsky, Rochester, & Shannon, 1955), which was held one year later in summer 1956. This workshop is generally seen as the official birth date of modern AI research. A huge AI hype followed, after the first “intelligent” systems were developed (e. g. ELIZA, an interactive program, which was able to carry on a dialogue in English language on any topic (Joseph Weizenbaum, 1966)), and huge promises were made by prominent AI researchers ("machines will be capable, within twenty years, of doing any work a man can do." - Herbert Simon, 1965 (Goble, 2012), “In from three to eight years we will have a machine with the general intelligence of an average human being.” - Marvin Minsky, 1970 (Darrach, 1970)). At that time, the first expert system was developed as well, automating the decision-making process and problem solving behavior of organic chemists by using detailed, task-specific knowledge as a source of heuristics and seeking generality through automating the acquisition of such knowledge (Lindsay, Buchanan, Feigenbaum, & Lederberg, 1993).

However, after James Lighthill (1973) reports to the British Science Research Council that “in no part of the [AI research] field have discoveries made so far produced the major impact that was then promised”, governmental support for AI research was reduced dramatically. What followed, is generally called the “AI winter”, describing this period of time with disappointed and lost public interest in the field and very limited funding. However, over the years, new approaches occurred. Probabilistic methods were found to function better than rule-based systems, for example in the area of language translation (Brown et al., 1988). In 1997, the IBM computer Deep Blue beats a reigning world chess champion for the first time. Similarly, the quiz show “Jeopardy!“ was won by an AI in 2011 using a huge knowledge base, and an AI based on artificial neural networks and tree search defeated the best human Go player in 2016. Modern AI research has many facets, ranging from understanding human speech, to competing at a high level in strategic games, to controlling autonomous cars.

Figure 1.4: Example of an autonomous car controlled by an artificial intelligence. Image source: https://www.digitaltrends.com/cars/chicago-may-ban-autonomous-cars/

As can be seen by this brief historical overview, and as is concluded by (Russell, S. & Norvig, 1995), too, the field of artificial intelligence research underwent several cycles of success and cutbacks, and new methodological approaches were developed constantly, how to create intelligent machines. Interestingly, the border of what can be seen as being “intelligent“ was moved further and further, never being reached. Only a short time after a task was accomplished by an artificial intelligence, which was previously being stated to require intelligence to solve, this point of view was changed, creating the so called AI effect (McCorduck, 2004). Until today, all forms of AI have rather been experts in a specific field or kind of task (e. g. playing chess) – the so called weak form of AI. The opposite, a strong AI displaying all forms of human intelligence (similar: artificial general intelligence) has not been created, and a valid question is, how such a form of intelligence could be tested (Goertzel, 2014). A number of tests have been proposed for this, with the so called Turing test being one of the most prominent ones. In brief, it tests, whether an AI is able to show intelligent behavior, which is indistinguishable from human behavior, by letting human interrogators ask written questions simultaneously to a human and an AI, trying to figure out from the answers, who is who (Turing, 1950).

1.5.2 Machine learning and deep neural networks

Recent AI research has seen a lot of progress in the field of machine learning. This subfield of AI describes the computer’s ability to learn without being explicitly programmed (Samuel, 1959). Even though other promising machine learning approaches exist, for example the decision tree-based random forests (Breiman, 2001), one extremely powerful technology are the so called deep neural networks (Lin, Tegmark, & Rolnick, 2017).

Artificial neural networks (for a detailed, formal introduction and a broader overview of the topic, see Schmidhuber (2015), Yegnanarayana (2009) or Goodfellow et al. (2016)) are based on the idea of modeling natural neural networks in computers – an approach, which foundations were laid by McCulloch & Pitts (1943). Like their natural counterparts, these digital neuron networks can receive inputs, transform the input and generate outputs. The transformation of the input takes place both in the neurons themselves (e. g. by summing up all inputs and applying a mathematical activation function to it in order to obtain an output value) and due to the complex, non-linear interplay between the neurons, with the output of one neuron being used as input to other neurons.

The strength of this approach arises, when the artificial neural network is modified according to certain rules – a process called learning. The goal of this is to let the neural network learn to generalize over different inputs, in other words, to be able to create meaningful output values also for before unseen data. Supervised and unsupervised learning are two categories of this, with supervised learning describing the idea of presenting input-output pairs to the network, making the network learn to calculate a specific output for given inputs, e. g. for pattern recognition purposes. In the case of unsupervised learning, the network is only presented inputs, and tries to describe a hidden structure in the input, e. g. finding clusters in the data.

The first idea, how a learning rule for modifying a network during the learning process could look like, was introduced by (Hebb, 1949), suggesting that “cells that fire together, wire together”, i. e. if neuron A in a network often causes neuron B having a high output value, the connection between them is strengthened. Important other learning algorithms like the backpropagation algorithm followed (Rumelhart, Hinton, & Williams, 1986), enabling the neural networks to learn more complex functions than was possible with the Hebbian learning rule.

Organizing the learning process is commonly subdivided into two parts: training and testing. During training, the network is presented a subset of the available data (training data set), and for each data point, an error function compares the network’s output with the target output values. Based on these deviations, the modifications of the network are applied. Simultaneously, a second subset of the available data (validation data set) is used in order to validate the neural network’s performance on unseen data, which helps prevent the so called *overfitting*, i. e. preventing the neural network from learning only the exact training examples, but not generalizing over them. Finally during testing, after the training and validation process resulted in a sufficiently good performance of the neural network on the validation data set, it is tested on the third part of the available data (the test data set), in order to estimate the neural network’s ability to handle completely new, unseen data.

Various types of neuron models, activation functions, learning rules, error functions and network architectures have been developed, resulting in more and more powerful artificial neural networks, which are capable of solving a manifold of problems[12]. Moreover, several

12 For the sake of brevity and since these details are not of the greatest importance for the rest of this dissertation, the interested reader is referred to the literature cited at the beginning of the chapter for further information.

open source frameworks such as Keras (Chollet, 2015) or tensorflow (Abadi, Barham, et al., 2016) exist for developing. Additionally, due to the increasing computational power of nowadays computers, and due to further theoretical improvements in the field, extremely large artificial neural networks can now be constructed and trained successfully (LeCun et al., 2015). This lead to the phrase “deep learning”, as the neural networks can consist of dozens or even hundreds of layers. Moreover, massive amounts of training data for many different problems is available nowadays due to the internet, smart sensors and devices, increasing the performance of neural networks further, succeeding human performance in many fields. One example for this is the success of deep neural networks at image classification, exceeding human performance in describing what can be seen in a picture (He, Zhang, Ren, & Sun, 2016).

Figure 1.5: This figure from (He et al., 2016) depicts image labels generated by an artificial neural network (numbered 1 to 5) as well as the "ground truth" (GT), i. e. the official label of the image from the image recognition competition.

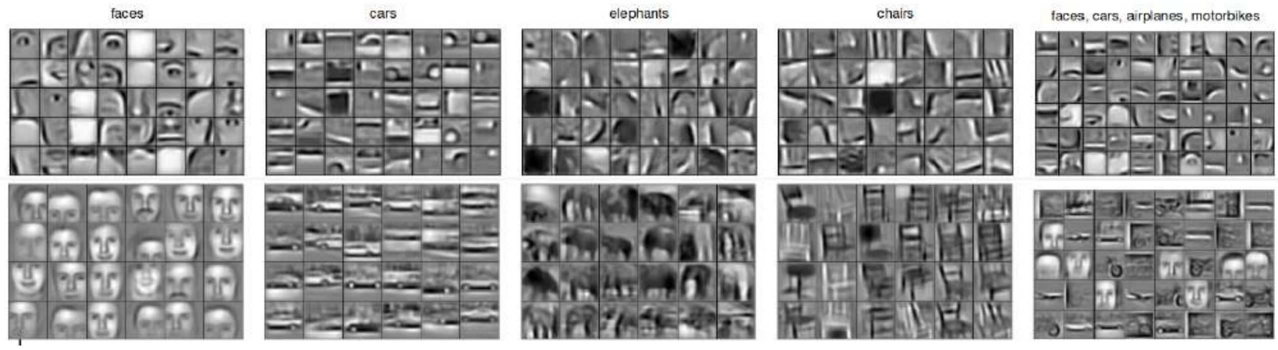


Figure 1.6: This figure from (Lee, Grosse, Ranganath, & Ng, 2009) depicts the features learned by so called convolutional neural networks, which have been trained on images of faces, cars, elephants, chairs, and a combination of faces, cars, airplanes and motorbikes, respectively. The first row shows the features learned by the neurons of an early layer of the network, the second row those of a later layer. It can be seen, that convolutional neural networks can learn hierarchical representations of images, i. e. identify common local patterns in the data (e. g. eyes or noses in the case of faces), as well as combinations of these, in order to classify before unseen pictures into one of the categories.

Figure 1.7: This example of a picture (right) created by an artificial neural network based on two input images (left) illustrates, that by understanding and combining the conceptual elements of a data structure (in this case: an image), new creative data representations can be formed. Image source: https://www.boredpanda.com/inceptionism-neural-network-deep-dream-art/?media_id=419452

1.6 Sleep and dream research and artificial intelligence

Several ways can be thought of, how artificial intelligence and machine learning can be used in the field of sleep and dream research. And also vice versa, AI can benefit from the knowledge created by sleep and dream researchers. In fact, research in both directions has been conducted already, as the following non conclusive list shows:

- By studying the nightly brain, mechanisms can be identified, which could lead to significant improvement of artificial neural networks, e. g. as used for unsupervised neural networks in (Hinton, Dayan, Frey, & Neal, 1995) or, regarding the function of sleep for memory consolidation processes or for restructuring knowledge, by (Walker & Russo, 2004). In these cases, artificial neural networks were trained in a similar way like natural neural networks learn, increasing their performance.
- Expert systems have been developed, which aid physicians in diagnosing sleep disorders (Ohayon, Guilleminault, Zulley, Palombini, & Raab, 1999), as well as astronauts in recording sleep data (Callini, Essig, Heher, & Young, 2000).
- Numerous approaches to AI-driven sleep stage classification using machine learning exist, which use as different AI methods such as linear discriminant analysis (Fell, Röschke, Mann, & Schäffner, 1996; Kayikcioglu, Maleki, & Eroglu, 2015), artificial neural networks (Hsu, Yang, Wang, & Hsu, 2013; Oropesa, Cycon, & Jobert, 1999; Park, Pa, & Jmn, 2000; Robert, Guilpin, & Limoge, 1998), support vector machines (Crisler, Morrissey, Anch, & Barnett, 2008; Koley & Dey, 2012; Lajnef et al., 2015), or genetic algorithms combined with a rule-based expert system (Kim & Park, 2000).
- AI methods were also found useful for automatically detecting microstructures in the sleep recordings, e. g. sleep spindles (Schimicek, Zeitlhofer, Anderer, & Saletu, 1994; Sinha, 2008; Ventouras et al., 2012; Wallant, Maquet, & Phillips, 2016).
- Furthermore, even approaches to measure the sleep quality in real-time based on AI-driven analysis of the ECG signal using support vector machines exist (Bsoul, Minn, Nourani, Gupta, & Tamil, 2010).
- Lastly, AI-driven approaches are also used for sleep medical purposes, e. g. for detecting sleep apnea in ECG sleep recordings (Babaeizadeh, White, Pittman, & Zhou, 2010).

No attempt could be found to automate complex sleep laboratory studies in the field, in order to enable naive crowd subjects to conduct them on their own in the field. Most likely, this is due to the unavailability of suited polysomnographic recording devices. As seen in chapter *1.4.2 Field studies using electronic tools*, tools which give access to raw data, which are cheap enough for crowd experiments, and which have been shown to have a good polysomnographic data quality are not available on the market. Logically, as long as there is no suited hardware, it makes no sense to develop specialized software, which enables naive crowd subjects to conduct polysomnographic experiments on their own and which uses a machine learning based approach for real-time sleep staging and other experiment control.

In order to change this, two things need to be developed for crowd-based polysomnographic sleep and dream experiments:

- a polysomnographic recording device, which is cheap, easy to use, has real-time raw data access, and which makes interactive sleep experiments in a sleep laboratory fashion possible at the subject's home,
- a piece of software, which enables naive crowd subjects to conduct the sleep and dream experiments on their own at home, i. e. which instructs the subject and which takes over the experiment control when the subject is asleep using state-of-the-art machine learning algorithms.

1.7 Aim of this dissertation

The aim of this dissertation is to develop and to evaluate such a portable sleep and dream experiment system consisting of hardware and software, which enables a crowd of naive subjects to conduct complex polysomnographic experiments at home. In other words, the goal is to make automated, interactive, crowd-based polysomnographic sleep and dream experiments possible. The name of this new system will be the *Traumschreiber* system (from German words "Traum" (dream) and "Schreiber" (writer)).

In order to test the suitability of the Traumschreiber system for complex crowd-based automated polysomnographic sleep and dream experiments, it will be evaluated regarding its data quality in comparison to sleep laboratory polysomnography, its usability, its production process and production costs, and its overall applicability for this type of studies.

2 Methods

2.1 Overview

An iteration based approach was chosen for developing the Traumschreiber system. This means that the system was repeatedly improved and tested, until the point was reach, at which a prototype was ready to be evaluated in a sleep laboratory and in field studies. In this chapter, the methods of the development process are described first, and afterwards the methods of the evaluation of the Traumschreiber system are presented.

2.2 Development

2.2.1 Defining the requirements towards the Traumschreiber system

Based on the experiences from own sleep laboratory and field studies, it was first defined, what the new Traumschreiber system had be able to do in order to fulfill its main purpose – i. e. enabling complex crowd-based automated sleep and dream experiments, including polysomnographic recording of EEG, EOG, EMG and ECG with good data quality, and real-time stimulation of the sleep. The requirements were regularly updated based on the testing during development, and based on the feedback obtained at conferences or of researchers, who tried out the Traumschreiber system (see chapter *2.2.2 Iteration-based development*).

The Traumschreiber system was designed to consist of

- a Traumschreiber sleep mask, which is worn by the experiment subject during the night, and which records and stimulates the subject's sleep,
- a Traumschreiber experiment station, which is placed next to the subject's bed, and which assists the subject, e. g. gives instructions and analyzes the recorded data in real-time,
- a Traumschreiber data analysis software package, which enables the researcher to perform the most important analysis steps in a standardized way after the data are recorded. Further details are depicted in the figures 2.1, 2.2 and 2.3[13] for each part.

13 The mind maps were created using the open-source WiseMapping tool.

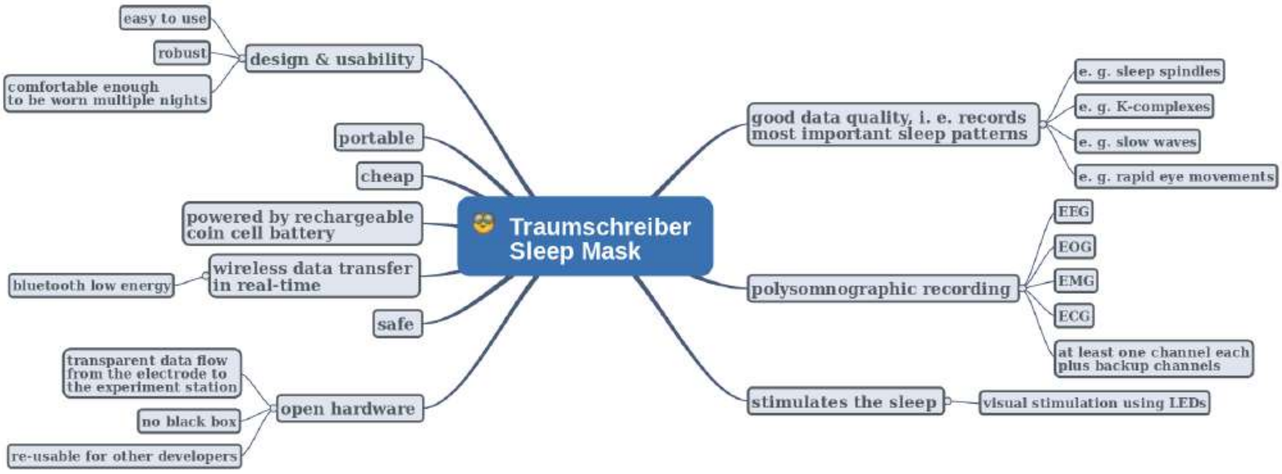


Figure 2.1: Requirements towards the Traumschreiber sleep mask.

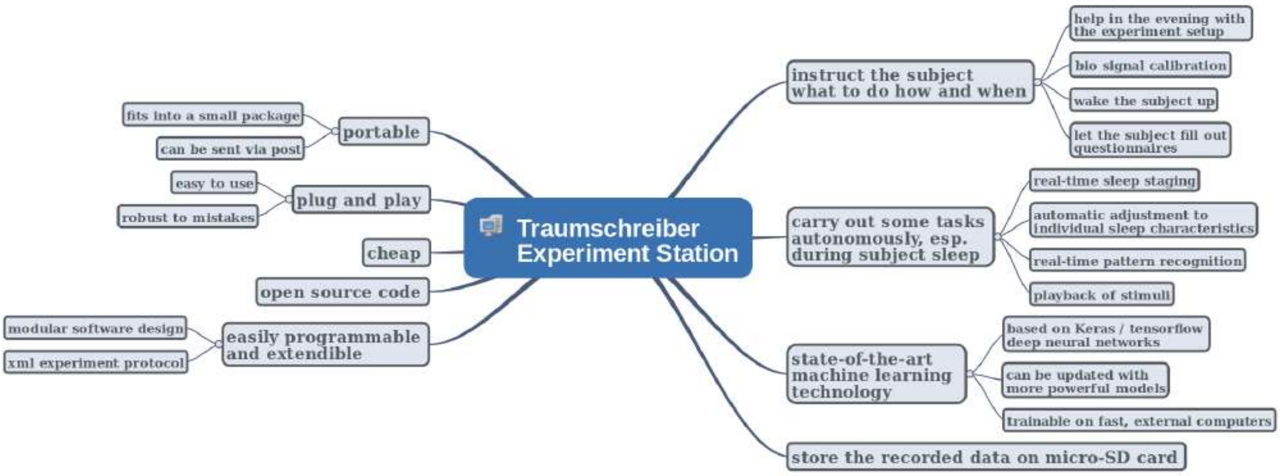


Figure 2.2: Requirements towards the Traumschreiber experiment station.

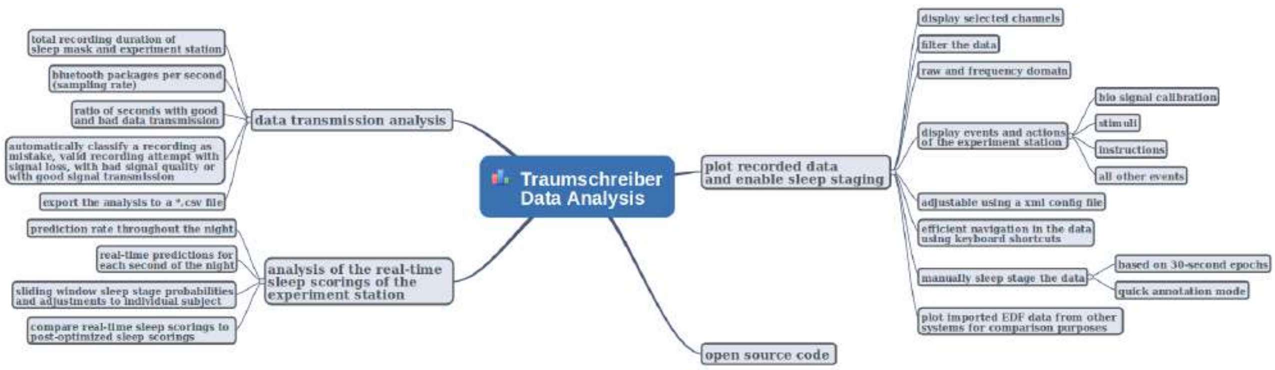


Figure 2.3: Requirements towards the Traumschreiber data analysis package.

2.2.2 Iteration-based development

During development, several decisions had to be made, how to best achieve the features depicted above. Should multi-usable conductive cloth be used as an electrode material, single use sticky electrodes, cup electrodes, a standard EEG cap, or some sort of a plastic helmet? How many channels are necessary for recording sleep physiological data; should there be backup channels in case some channels have a bad signal? How shall the data be transferred between the sleep recording device and the experiment station: via wi-fi, bluetooth, or bluetooth low energy? Should the experiment station rather be a Raspberry Pi minicomputer, a smartphone, or directly integrated into the sleep mask? Should acoustic stimuli and instructions be delivered to the subject via speakers or headphones? Should the recorded data be processed locally, or directly transferred to a server via internet? In what kind of format should the data be stored? Which programming language should be used for programming the experiment station? Which frameworks are suited best for implementing the artificial intelligence part of the system? Not even mentioning all the electric engineering decisions, e. g. regarding the hardware filters, the placement of the components on the PCB, the selection of the optimal microcontroller considering costs, energy consumption, availability, price, and of course technical properties (speed, ADC, connection types, etc.), and many more – in the end, a complete EEG was constructed from scratch. And where should this be produced (China, Germany)? How should it be financed?

These questions illustrate well, how many different decisions had to be made during development, ranging from textile design, to electric engineering, to machine learning, to sleep and dream research. Trying to answer all questions at once right in the beginning of the project was not possible. Instead, an iteration-based approach was chosen.

Each iteration consisted of four main steps:

1. Defining / adjusting the requirements towards the new version of the system, and setting priorities. On the one hand, this was based on the experiences gained from testing the previous version in self-experiments. On the other hand, input from other sleep researchers and clinicians was obtained. Some of these researchers tried out prototypes of the system, others gave feedback at oral or poster presentations about the system at scientific conferences[14]. Furthermore, interested hobbyists

14 23rd Annual Meeting of the German Sleep Society 2015 - oral presentation at the young scientists

(mostly lucid dreaming enthusiasts experienced with other consumer devices such as the Zeo) suggested features to be included in the next development cycle. Last, but not least, the experiences of the author of this dissertation from the sleep experiments at the Osnabrueck University sleep laboratory and at the Max Planck Institute of Psychiatry, Munich, inspired the development of new features, and helped setting the priorities.

2. Finding ways, how to implement the needed features, or how to improve the existing solution. This depended strongly on the type of task, which was conducted, and was done via internet research (e. g. finding new electric components), trial and error (e. g. trying out different approaches to make the sleep mask more comfortable), measuring and electrical engineering using an oscilloscope, by-hand calculations and specialized software (e. g. designing hardware filters, drafting the electrical circuit diagrams and PCB layouts), programming and refactoring (e. g. new software features or the microcontroller code), and group discussions (e. g. which tasks should have priorities or whether to choose a more expensive but better electrical component or not). Furthermore, the autodidactic learning part should not be underestimated, as no professional electrical engineer or product designer was involved in the development, and most skills were obtained in a self-taught way.

Figure 2.4: Picture of the technical development process (Johannes Leugering)

workshop: “The Traumschreiber Project”; Hofgeismar Young Scientists Workshop of the German Sleep Society 2016 - oral presentation: “Updates on the Traumschreiber Project”; MetaRheinMainChaosDays 2016 of the Chaos Computer Club - oral presentation: “Traumschreiber: Eine Open Source Schlafmaske, die Hirnaktivitäten messen und beeinflussen kann”; 23rd Congress of the European Sleep Research Society 2016 - poster presentation P214: “‘Traumschreiber’: measuring and manipulating human sleep with a portable high-quality but low-cost polysomnographic system”.

3. Obtaining finances for the production of the new prototype. In the beginning, the Neuroinformatics department of Osnabrueck University funded the project, as well as the Institute of Cognitive Science at Osnabrueck University. In later development stages, when the production became more and more expensive, also external financial sources had to be acquired, e. g. the Hans-Mühlenhoff-Stiftung Osnabrueck (2,000 EUR) and the German Sleep Society (5,000 EUR).
4. Producing a new prototype based on the current iteration's developments, and testing it. In the very beginning of the project, the electric parts of the sleep recording device were soldered on a breadboard and directly tested. In subsequent development stages, the printed circuit board was designed in Osnabrueck, produced in a factory in Germany, and assembled with the electric components in Osnabrueck, as well as tested in Osnabrueck. During the last three developments iterations, the printed circuit board was designed in Osnabrueck, produced and assembled at much lower costs in China, then imported to Germany, and tested in Osnabrueck. The textile parts were completely self-designed in the very beginning. Later, ready-to-use sleep masks were bought in Germany and processed further in Osnabrueck, and finally much cheaper (and at the same time even more comfortable) sleep masks were imported from China and processed further in Osnabrueck, i. e. working the electric parts into the sleep masks. Testing consisted of a) wearing the device during self-recordings for several nights and then analyzing the data quality by plotting and visual inspection, b) analyzing the comfort of the sleep recording device, and c) analyzing the necessary work steps for an experimenter to conduct an experiment. If the new features of the system worked well, they were kept for the next development iteration, if not, they were either changed or removed in the next development round.

During all the above mentioned work steps, external advice or work force was acquired – in some development iterations for the one step, in other iterations for another step. For the textile development, for example, a cooperation with the study program “Textile design” at Osnabrueck University was established, and cognitive science and textile design students drafted and produced several textile designs, how high-tech sleep masks should look like. Moreover, some work packages have been “outsourced” to bachelor's theses of cognitive science students, supervised by the author of this dissertation (all attached in the online appendix): a literature review of sleep stimulation experiments (Braumann, 2016), testing the EOG capabilities of the sleep recording device (Laubisch, 2016), conducting a sleep laboratory validation (Nienhaus, 2017), testing the usability and functionality of the system (Mandt, 2017), and programming a first version of an Android app (Jäkel, 2017).

#	Features, comments	Date	Production place
v0.0	Breadboard based EOG measurement with oscilloscope, inspired by OpenEEG project	07/2014 to 09/2014	At home
v0.1	First PCB, measures EOG and transmits the signal via cable to computer	10/2014 to 07/2015	PCB: produced in German factory, equipped by hand in Osnabrueck
v0.2	Measures sleep EEG, transmits the signal via bluetooth 2.0, has LEDs on external board, which can be controlled via bluetooth, limited auditory stimulation via headphone plug onboard	08/2015 to 11/2015	PCB produced in German factory, equipped by hand in Osnabrueck
v1.0	Uses self-designed signal amplifiers (op-amps instead of IAs), LED board split up into driver board and satellite LED boards, used during textile design seminar	12/2015 to 03/2016	PCB produced in German factory, equipped by hand in Osnabrueck
v2.2	BLE implemented (but does not work), eight channels instead of four, uses IAs again for signal amplification, has flat connectors for the cables, first field study	04/2016 to 12/2016	PCB produced and equipped in China
v2.4	BLE works, LED onboard, battery onboard, DIN cable connectors, sleep laboratory validation, most boards faulty due to bad production quality of new Chinese manufacturer	12/2016 to 03/2017	PCB produced and equipped in China
v2.5	Slightly changed LEDs, ordered at first Chinese manufacturer again (good production quality) **Final version**	03/2017 to 08/2017	PCB produced and equipped in China

Table 2.1: Overview of the main electric development iterations

In total, seven development iterations were conducted. Especially obtaining the finances (by applying for small grants), university and customs bureaucracy, production abroad and testing in experiments were very time consuming. The final version of the system (including the sleep recording device), as well as the exact work steps, how to produce it, are described in chapter *3.1 Development* and *3.2.3 Third study: Preparing polysomnographic crowd experiments*.

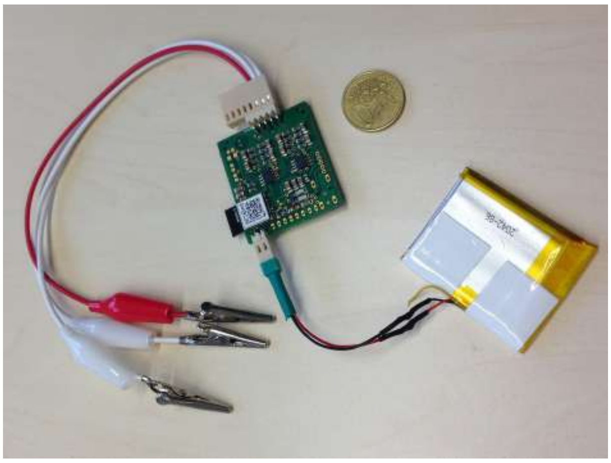


Figure 2.5: Electric prototype version 0.2. It had to be connected to an external battery, it had only four data channels and used bluetooth 2.0 instead of BLE for data transmission.

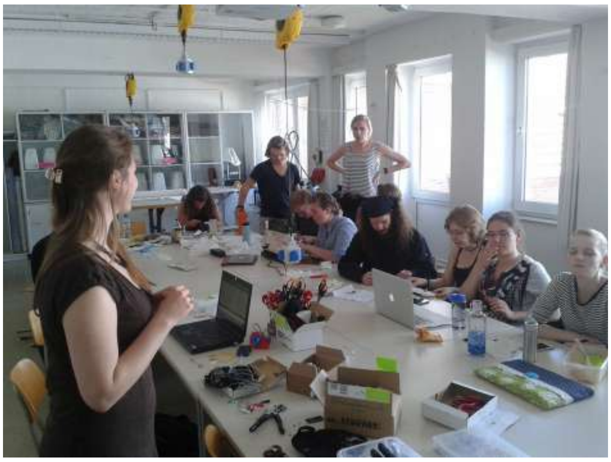


Figure 2.6: Picture of the textile design seminar.

Figure 2.7: One intermediate result of the textile development process.

Figure 2.8: The first self-made prototype of the textile part of the sleep mask. Note that it is flat, i. e. moving the (closed) eyes under this sleep mask is uncomfortable.

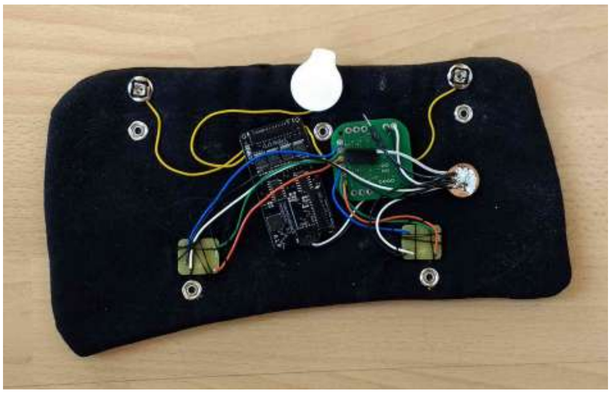


Figure 2.9: Traumschreiber sleep mask version 1.0. Note the several boards for measuring sleep, for LED control and the actual LEDs, all connected via cables.

Figure 2.10: Traumschreiber sleep mask 1.0 together with the eye cover. Note that the light of the right LED shines through the eyelet.

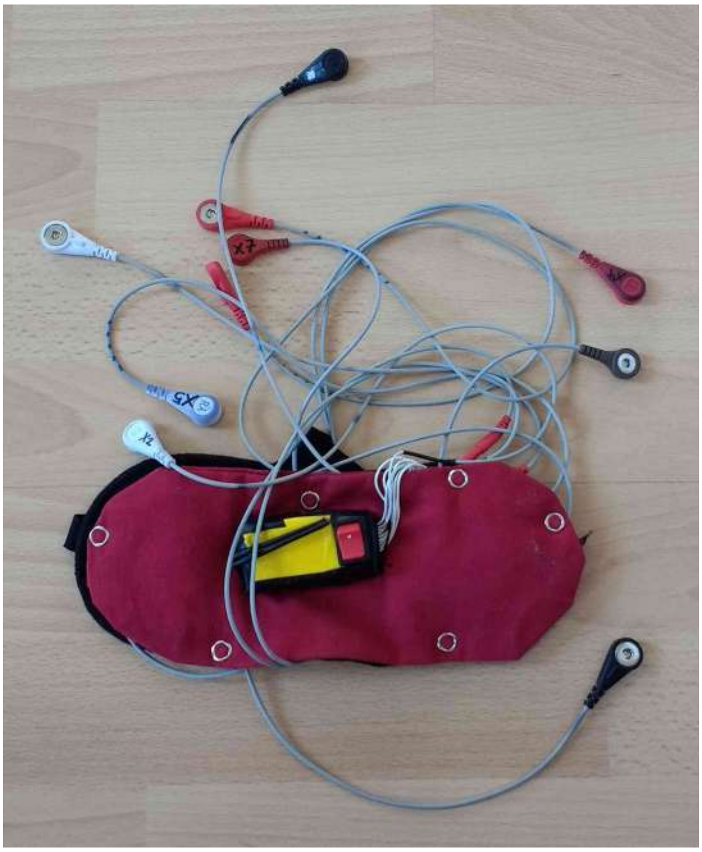


Figure 2.11: Traumschreiber sleep mask version 2.2. Note that the cables are far too long (making the sleep mask heavy and uncomfortable) and that the electronics are not hidden inside the mask.

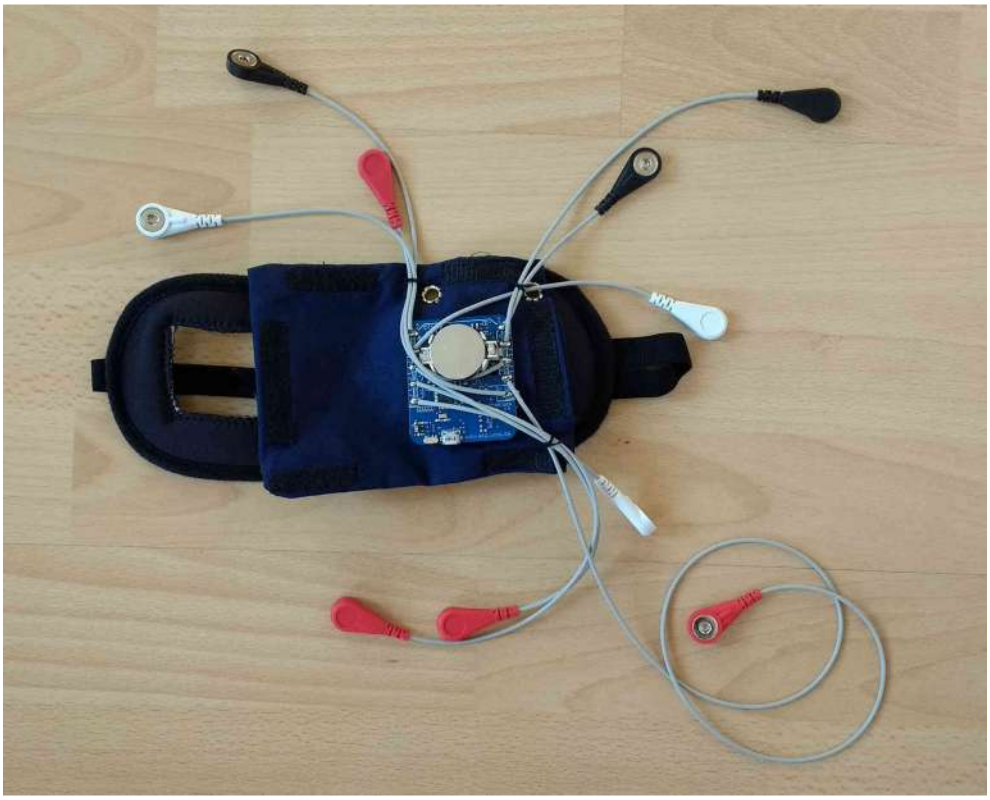


Figure 2.12: The final version of the Traumschreiber sleep mask (open pocket).

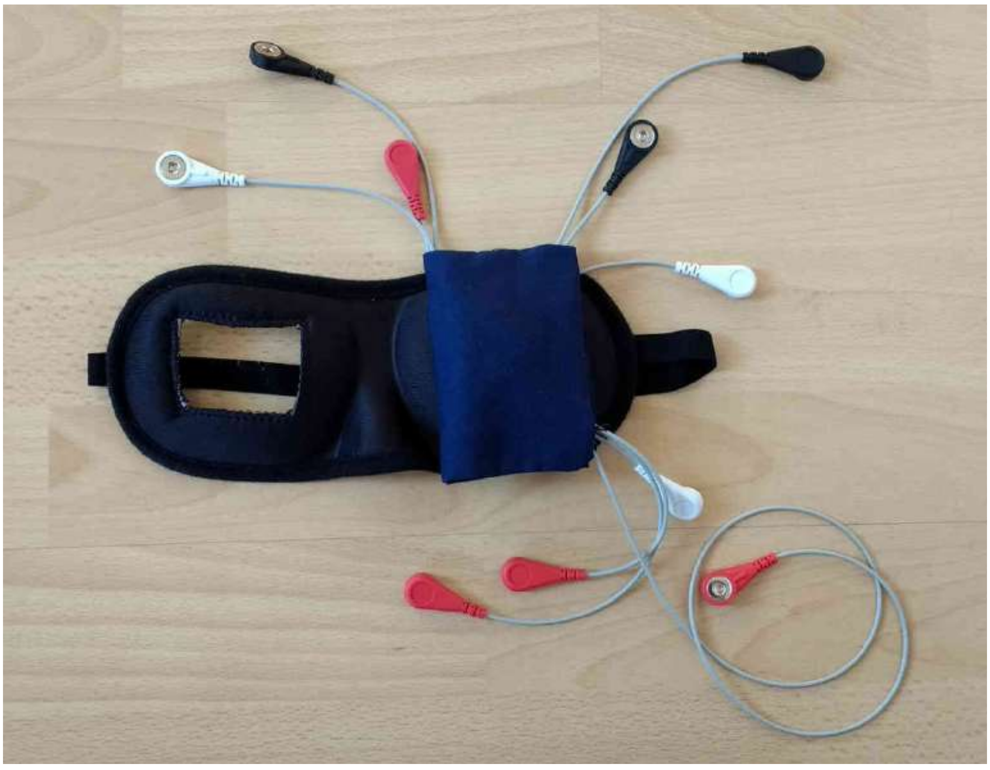


Figure 2.13: The final version of the Traumschreiber sleep mask (closed pocket).

2.3 Evaluation

After developing the system, it was evaluated in several studies. The purpose of this was to analyze, how well the system is suited for real-world usage of sleep researchers. It is suggested to first read the description of the development results (chapter *3.1 Development*) before continuing with reading this chapter.

2.3.1 First study: Sleep laboratory validation

2.3.1.1 Aim

Can the same sleep characteristics be measured using the Traumschreiber system and a commercially available, worldwide used medical polysomnographic system? What are the differences in the recorded data? In order to answer these questions, the Traumschreiber system was validated against such a commercial medical polysomnographic system in the sleep laboratory of Osnabrueck University.[15]

2.3.1.2 Subjects

Seven healthy students (2 female, age between 18 and 27 years) were recruited via email announcement. Exclusion criteria were: younger than 18 years; pregnancy or breastfeeing; suffering from psychic or psychosomatic illnesses, neurologic diseases, hallucinations, or sleep disorders (evaluated using the standardized LISST questionnaire (Schürmann et al., 2001)); taking sleep drugs, psychotropic drugs, pain killers, or antihistaminics. The subjects were asked to not consume any alcohol within the last 24 hours before the experiment, to not nap during the day, and to sleep regularly the night before the experiment. The subjects spent one night each at the Osnabrueck University sleep laboratory. They signed written informed consent and were compensated with 3 subject hours, which they needed for their degree. One subject had to be excluded from data analysis, as the battery of the sleep mask was not charged and recorded only a part of the night.

2.3.1.3 Materials and procedure

Using an ALICE 3.5 commercial medical polysomnographic system, chin EMG, EOG, frontal, central and occipital EEG (2 channels each) as well as ECG were recorded using

15 The study was conducted together with Frederik Nienhaus, who wrote his bachelor’s thesis about this experiment (Nienhaus, 2017, supervised by the author of this dissertation).

gold cup electrodes (central and occipital EEG) and single use sticky electrodes (Covidien Kendall H135SG) for the other recording locations. Simultaneously, the prototype version 2.4 of the Traumschreiber system was used, which did not differ from the final version of the Traumschreiber system regarding the relevant technical specifications of the PCB or its components (same filters, same microprocessor, same bluetooth chip, etc.). The electrodes of the Traumschreiber sleep mask were of the same type as the ALICE system's and were attached directly next to them, with the intention to record similar physiological activity. An exception were the central and occipital EEG electrodes of the ALICE system: Since only single use sticky electrodes are used in the standard Traumschreiber recording setup, which do not hold in hair, no Traumschreiber electrodes were attached near them. Thus, in total 14 electrodes were fixed in the face of the subjects, plus additional 11 electrodes behind the ears (references), in the hair (EEG), and on the chest (ECG) – see figure 2.14.

The subjects were sent to bed around 11 PM, and were woken up around 7 AM, resulting in a total of 8 hours of bed time.

2.3.1.4 Data analysis

Figure 2.14: Sticky electrodes of both systems were placed on the forehead, recording EOG and frontal EEG.

30-second epochs of the data were plotted in 15 minute intervals using the post_visualization.py script of the Traumschreiber data analysis software package and a modified EDF-compatible version of it for the ALICE data. This means that a snapshot of the recorded data was produced for each of the time points 11:00 PM, 11:15 PM, 11:30 PM throughout the whole night to 6:30 AM, 6:45 AM, 7:00 AM for both systems including all channels. Both systems were synchronized in time using the bio signal calibration markers, i. e. the calibration procedure at the beginning of each night, for example clenching the teeth or moving the eyes in a systematic manner.

The plots of the data of the Traumschreiber system and of the ALICE system were then compared by visual inspection. As a first step, each plot of the Traumschreiber system was analyzed, whether there were clearly recognizable slow waves, sleep spindles, K-complexes, a low amplitude and mixed frequency EEG signal, rapid eye movements, slow eye movements, eye blinks, clearly visible heart beats in the ECG signal, arousals, alpha waves, or vertex waves. Next, the same was done for the plots of the ALICE system. It was then denoted, whether the two systems showed similar elements in each corresponding epoch, or whether some graphoelements could only be seen in one of the recordings. Furthermore, the differences were described in detail by words, if necessary, e. g. if sleep spindles were not detectable in the Traumschreiber system's plot, because they were of central origin. As a second step, every epoch of each system was classified into the categories wake, REM, N2, N3 and ambivalent sleep. It is sometimes impossible to rate, whether a single 30-second data bin depicts wake, N1, N2 or (tonic) REM sleep, because this depends on the classifications of previous epochs, too. For these cases, the *ambivalent* sleep category was used. The classifications of both systems were then compared. Descriptive statistics were used to summarize the results.

2.3.2 Second study: Home-based prototype study

2.3.2.1 Aim

In this study, a prototype (version 2.2) of the Traumschreiber system was tested for the first time at the subjects' homes. The aims were

- to investigate, whether the crowd subjects were able to conduct the study with the Traumschreiber system as intended,
- to check the data quality of the recordings,
- to analyze the effect of the system on the sleep and dreams of the subjects,
- to find out, how easy to use and comfortable the system was for the subjects,
- to determine, how the system should be further developed, and
- to conduct a small scientific study as a side experiment with the new system, namely, to analyze the effect of acoustic stimulation from the Traumschreiber system on sleep and especially on the occurrence of arousals.[16]

2.3.2.2 Subjects

24 healthy students (15 female, age between 18 and 27 years) were recruited via word-

16 The study was carried out together with Laura Sophie Mandt, how wrote her bachelor thesis about this experiment (Mandt, 2017, supervised by the author of this dissertation).

of-mouth advertising and e-mail announcements. Exclusion criteria were: younger than 18 years; pregnancy or breastfeeing; suffering from psychic or psychosomatic illnesses, neurologic diseases, hallucinations, or sleep disorders (evaluated using the standardized LISST questionnaire (Schürmann et al., 2001)); taking sleep drugs, psychotropic drugs, pain killers, or antihistaminics. The subjects were asked to not consume any alcohol within the last 24 hours before the experiment, to not nap during the day, and to sleep regularly the night before the experiment. One subject did not meet these criteria and was thus excluded from further data analyses. The subjects received no financial payment, but three subject hours needed for their degree. All subjects signed written informed consent.

2.3.2.3 Materials

The prototype Traumschreiber system, which was handed out to the subjects of this study, consisted of the sleep recording device, a Raspberry Pi 3 Model B minicomputer with a touchscreen, two small USB speakers “Trust Leto 2.0 USB”, and a sufficient amount of single use sticky electrodes (Covidien Kendall H135SG). The sleep recording device was a sleep mask similar to the final development version 2.5, but with a larger battery pack (three AAA batteries instead of a coin cell battery) and much longer cables (90 cm each), which were connected to the board via a connector and not directly soldered to it (see figure 2.11). The recording locations used in this experiment differed slightly from the suggested electrode positions in the final version of the system (two ECG channels instead of one, only two dedicated EEG channels (Fp1-F1, F1-F2), no reference electrode behind the ear). Moreover, the on-off-switch of the sleep mask was larger and on the outside of the cloth, and the sleep mask textile was larger and had a different color, but was of the same kind of fabric as the final version. The relevant electric parts of the prototype did not differ from the final version of the system regarding the technical specifications of the PCB or its components (same filters, same microprocessor), however, the data were transmitted using an older bluetooth 2.0 chip.

The subjects were instructed verbally and via a written text how to conduct the experiment, and received a demonstration video showing how to attach the electrodes (see figure 2.15). They were also asked to fill out questionnaires during the whole experiment.

2.3.2.4 Procedure

Figure 2.15: Screenshot of the instruction video (Laura Mandt).

The subjects received the materials from the experimenter during the day, and were instructed to record the following night with the Traumschreiber system, i. e.

- to watch the instruction video,
- to switch on the minicomputer and to place the speakers in 2m distance from the pillow,
- to follow the minicomputers instructions, (put on the electrodes at the correct positions, connect them to the sleep mask, switch on the sleep mask, go to bed),
- to conduct the bio signal calibration (e. g. move the eyes in a specific way, clench teeth) according to the instructions of the minicomputer,
- to try to sleep normally,
- to disconnect and switch off everything the following morning,
- to fill out questionnaires about the past night's sleep (Schlaf-Fragebogen A (Görtelmeyer, 1986)) and about the Traumschreiber system (self developed questionnaire "Fragebogen zum Gebrauch des Traumschreibers", see online appendix), as well as a dream report, if they could remember any dream from the experimental night, and
- to return everything the same day.

A control night was conducted one week later, but without the Traumschreiber system, i. e. only questionnaires had to be filled out by the subjects.

The batteries of the sleep recording device were re-charged by the experimenter in between the recordings of the different subjects.

2.3.2.5 Arousal threshold experiment

In order to investigate, how well the system was suited for conducting a small scientific experiment, a “toy” experiment was programmed into the system. The subjects were randomly assigned to one of two groups: one group (N=14) received acoustic stimulation from the system via its USB speakers throughout the whole first night in regular time intervals of 10 minutes, the other group (N=9) received no acoustic stimulation. The stimuli consisted of one second long 1000 Hz sine wave tones of random volume (ranging from not noticeable to room volume level). The goal was to analyze the effect of such a stimulation on the sleep quality, the feeling of recovery in the morning, as well as to confirm the hypothesis, that arousals are more frequent after louder stimuli than after quiet ones.

2.3.2.6 Data analysis

For investigating, whether the Traumschreiber system fulfilled its purpose at the subjects’ homes as intended, the different subfields of the experiment were viewed holistically. This included the following data analyses:

- The data quality of the recordings was assessed descriptively by analyzing the amount of time, during which data was transmitted from the sleep mask to the minicomputer, the sampling rate (number of data points per second), and by visually inspecting the signal quality in each channel for each recording.
- The effect of the system on the sleep of the subjects, as well as on dreaming, was assessed as follows: The Schlaf-Fragebogen A (SF-A) was evaluated as described in the manual by (Görtelmeyer, 1986) and yielded information about the total sleep time, the sleep quality and the feeling of recovery both in the first night and in the control night. For sleep quality, this was based on the time to fall asleep in the evening, the number of awakenings during the night, the length of the awake periods during the night, and a self-assessment of the subject, how constant, deep, restless, relaxed, undisturbed and good the sleep of the previous night was. For the feeling of recovery, this was based on the ampleness of sleep and whether, subjectively, the subject felt even-tempered, drowsy, cheerful, fresh, relaxed, and well-rested the next morning. The data of the first and the control night were then tested for statistically significant differences using Cohen’s effect size d (Cohen,

1992) and permutation tests ((Fisher, 1937), within-subjects), as well as the differences in the first night between the stimulation and the non-stimulation groups (between-subjects). Dreaming differences were assessed by analyzing the dream reports of the subjects regarding the number of recalled dreams, and the incorporation rates of the experiment and of the stimuli into the dream content, using descriptive and inferential statistics by calculating Cohen’s effect size d and permutation tests (between-subjects).

- For finding out, how easy to use and comfortable the system was for the subjects, descriptive statistical analyses were conducted based on the answers to the self developed questionnaire.
- The recorded data were visually inspected for each stimulation, in order to analyze the effect of acoustic stimulation carried out by the Traumschreiber system on sleep and especially on the occurrence of arousals. If the data quality was sufficient, the stimulus was classified as either eliciting no arousal (no change in frequency or amplitude after stimulus onset), or eliciting an arousal (distinct change in the signal visible after stimulus onset). Statistical analyses were then conducted in order to evaluate the stimulus loudness with respect to the classification of an arousal or no arousal applying Cohen’s effect size d, and checked for significance using permutation testing.
- Collecting ideas on how the Traumschreiber system should be further developed: This was conducted by summarizing the suggestions of the subjects from the questionnaires, as well as by reporting practical experiences with the system.

For all analyses, descriptive and inferential statistical data analyses were carried out, based on the questionnaires and the recorded data. The recorded data were plotted and analyzed using R and Python scripts (previous versions of the Python data analysis scripts of the final version of the Traumschreiber system).

2.3.3 Third study: Preparing polysomnographic crowd experiments

2.3.3.1 Aim

The goal of this analysis was to collect and describe all tasks, which have to be conducted in order to produce one or more entities of the new system. The preparation of study four (see chapter *2.3.4 Fourth study: A crowd-based polysomnographic experiment*) was analyzed for this purpose regarding production times and production costs. The knowledge obtained by this is especially important for the future usage of the Traumschreiber system in crowd-based projects, for which a beforehand calculation of production times and costs is needed. The description of how to prepare a Traumschreiber study starts with ordering the raw material and ends with handing out the ready-to-use experiment boxes to the subjects. This chapter can also be seen as a short tutorial for setting up a Traumschreiber system study.

2.3.3.2 Description of the preparation steps for a Traumschreiber system study

During the preparation of the fourth study of this dissertation, every preparation step was noted. This included both the hardware production and the necessary tasks for software setup. The tasks summary can be seen as rather independent of the total amount of Traumschreiber systems needed, as every system requires the same tasks to be done in order to obtain a working Traumschreiber system.

For determining realistic production times of each production task, the components of 20 Traumschreiber systems were ordered and put together by a person with amateur skills (at most!) regarding textile and electronic craftsmanship (e. g. sewing and soldering) - the author of this dissertation. The time needed for production was measured for each task. The production times of the work steps, for which a decrease in production time was expected due to learning and practical experience (e. g. soldering, sewing), were assessed for each sleep mask individually. The collected data were then summarized.

For calculating the costs per Traumschreiber system, the real costs for ordering the 20 Traumschreiber systems of the fourth study were noted. Additionally, it was described, whether the costs of the parts of the system stay rather constant regardless of the amount of systems being produced, or whether larger quantities lead to lower costs per system.

The 20 experiment boxes, which were produced in this study, contained the latest development version of the Traumschreiber system: a hightech sleep mask with electrode cables, a Raspberry Pi 3 Model B minicomputer with power adapter, a set of two USB speakers, a package of 50 electrodes, and the paperwork including subject information, consent form and questionnaires.

2.3.4 Fourth study: A crowd-based polysomnographic experiment

2.3.4.1 Aim

The main goal of the Traumschreiber system is to enable naive crowd subjects to conduct complex, interactive sleep and dream experiments on their own at home. The system assists the inexperienced subjects by instructing what to do when, and conducts major parts of the experiments automatically. This makes massively parallelized studies possible. In this fourth study, a demonstration experiment was conducted in order to test, whether the Traumschreiber system is capable of fulfilling its purpose, and if so, how well it is suited and where the pitfalls lie.

During three nights for each subject, a sleep experiment was conducted at the subjects' homes. One aim of recording multiple nights was to find out, whether the subjects became more practiced and secure in how to use the system after a few nights, and if this could be seen in the data transmission success or data quality. Furthermore, it was analyzed whether any first night effects could be found regarding sleep quality or comfort of the system. Another aim was to investigate, whether any signs of wear could be detected after three nights of use.

By recording more than 10 subjects simultaneously, the study also aimed at finding out more about the organizational demands for the scientist (before, during and after the experiment), and whether any problems can be expected because of this parallel design.

The experiment did not aim at delivering new scientific insights about sleep or dreaming in general, but should test the suitableness of the Traumschreiber system for complex automated sleep experiments at the subject's home - including automated real-time analyses of the recorded data. Moreover, the experiment aimed at finding out, how well the simplistic automatic adjustment of the sleep scoring algorithm to individual subject sleep characteristics functions.

An additional exemplary research question was chosen, which was not expected to reveal anything new about sleep or dreaming, but for which the results are mostly well known already. This way, the results of the here conducted experiment could be compared easily to previous studies, in order to see whether it is possible to obtain the same results with this new methodology as in previous studies with full polysomnography in the sleep laboratory. The research question was, in how far REM dreams differ from N3 sleep dreams, regarding how static and thought-like they are, how vivid they are, how bizarre they are, and how entertaining they are, and, additionally, how easily subjects can be woken up from the two sleep stages.

2.3.4.2 Subjects

18 healthy adults (13 female, age between 18 and 31 years) were recruited via the psychology students mailing list of the university. Exclusion criteria were: younger than 18 years; pregnancy or breastfeeding; suffering from psychic or psychosomatic illnesses, neurologic diseases, hallucinations, or sleep disorders (assessed using the standardized LISST questionnaire (Schürmann et al., 2001)); taking sleep drugs, psychotropic drugs, pain killers, or antihistaminics; and having a beard at the chin. The subjects were asked to not consume any alcohol within the last 24 hours before the experiment, to not nap during the day, and to be careful with the materials. Four subjects dropped out before the study started due to personal reasons (the study was conducted during the exam phase of the semester). Thus, 14 subjects collected one experiment box each. One further subject dropped out during the first night of the experiment due to a technical defect of the sleep mask (on-off switch broke), resulting in 13 subjects completing the experiment for three nights each. The subjects signed informed written consent prior to the study and received no payment, but six subject hours, which they needed for their degree.

2.3.4.3 Materials

Each subject received one personal experiment box as described in chapter *3.1.2 The components of the Traumschreiber system*, containing a sleep recording device, a Raspberry Pi 3 Model B minicomputer, which was connected via a USB sound card to two small USB speakers, a sufficient amount of single use sticky electrodes (Covidien Kendall H135SG), as well as printed out questionnaires for all three nights.

2.3.4.4 Time frame

The whole experiment took place within one week, from June 26th to July 2nd, 2017. The subjects were allowed to choose three out of seven nights on their own, during which they wanted to conduct the experiment.

2.3.4.5 Procedure

Before participation, the subjects were informed about the study's goal and methods, and signed informed written consent. Next, the subjects filled out the LISST questionnaire and the subject information sheet, which they then returned together with the consent form to the experimenter. If none of the exclusion criteria were met, the Traumschreiber box was handed out to the subject.

The subjects were instructed to record three out of seven nights with the Traumschreiber system on their own, i. e.

- to watch the instruction video of how to stick on the electrodes, how to put on the sleep mask, how to switch on the sleep mask, how to remove the electrodes and the sleep mask in the morning, and how to recharge the battery of the sleep mask (screenshots of the instruction video in figures 2.17 and 2.16, the whole instruction video can be found in the online appendix),
- to switch on the Traumschreiber experiment station and to place the speakers in 2m distance from the pillow,
- to follow the Traumschreiber experiment station's instructions, (stick on the electrodes at the correct positions, connect them to the sleep mask, switch on the sleep mask, go to bed),
- to conduct the bio signal calibration (e. g. move the eyes in a specific way, clench teeth, etc.) according to the instructions of the Traumschreiber experiment station,
- to try to sleep normally,
- to turn off the Traumschreiber experiment station by unplugging the power cable and to switch off the sleep mask in the morning,
- to recharge the sleep mask batteries by plugging the power adapter of the minicomputer into the sleep mask in the morning, and
- to fill out questionnaires about their sleep (Schlaf-Fragebogen A (Görtelmeyer, 1986)) and about the Traumschreiber system (self developed questionnaire "Fragebogen zum Gebrauch des Traumschreibers", see online appendix).

Additionally, the subjects were woken up up to six times in each experimental night, preferably during REM or N3 sleep (sleep staged automatically in real-time). In case no sleep stage was automatically detected, subjects were woken up after a maximum waiting time (see experiment xml file in figure 2.18, for details about the experiment xml file concept see chapter *3.1.2.2.1 Autonomous, easy to program experiments*). They then had to fill out a brief questionnaire about their dream, in case they had one prior to awakening, and were sent back to sleep again.

Figure 2.17: Screenshot of the instruction video (placing the electrodes).

Figure 2.16: Screenshot of the instruction video (connecting the electrodes to the sleep mask).

```xml
<root>
    <experiment>
        <sound>welcome.wav</sound>
        <sound>follow_evening_instructions_then_turn_on_TS.wav</sound>
        <wait_for>TS_connected</wait_for>
        <sound>connected.wav</sound>
        <sound>calibration.wav</sound>
        <sound>good_night.wav</sound>

        <wait_for>duration_00:10:00:000</wait_for>
        <wait_for>sleep OR duration_00:30:00:000</wait_for>
        <wait_for>duration_00:30:00:000</wait_for>
        <wait_for>REM OR N3 OR duration_00:45:00:000</wait_for>
        <arousal_procedure></arousal_procedure>
        <sound>night_questionnaire.wav</sound>
        <coding_number></coding_number>


                        ...


        <wait_for>duration_00:10:00:000</wait_for>
        <wait_for>sleep_onset OR duration_00:30:00:000</wait_for>
        <wait_for>duration_00:30:00:000</wait_for>
        <wait_for>REM OR N3 OR duration_00:45:00:000</wait_for>
        <arousal_procedure></arousal_procedure>
        <sound>night_questionnaire.wav</sound>
        <coding_number></coding_number>


        <wait_for>duration_99:00:00:000</wait_for>

    </experiment>
</root>
```

Figure 2.18: The experiment xml file of the fourth study. After the Traumschreiber experiment station played a welcome message, the subject was asked to follow the written and video instructions, and to switch on the Traumschreiber sleep mask, when lying in bed. Next, the subject was informed that the connection was established, and instructed to conduct the bio signal calibration. Next, the subject was wished a good night. Each arousal block consisted of waiting for sleep, detecting REM or N3 sleep, waking up the subject, asking to fill out the night questionnaire, and saying a random number, which had to be noted in the night questionnaire. The three dots stand for four further repetitions of the arousal block.

2.3.4.6 Data analysis

General results

First, the data of all subjects was evaluated, regarding whether the Traumschreiber system conducted its tasks as planned: Were the subjects instructed and enabled to record polysomnographic data on their own autonomously in their natural sleep environment? How reliable were the subjects? The answers to the questionnaires as well as the data of the recordings were summarized using descriptive statistics.

Data quality

The recorded nights were analyzed regarding the total recording duration and the transmission quality using the data_transmission_analysis.py script (see chapter *3.1.2.2.4 Data visualization and data analyses*). Based on the table and the plots created by the script, the nights were classified as “good transmission quality” (good signal quality determined by the script and full night recorded) or “bad transmission quality”. Descriptive statistics and inferential statistics based on Cohen’s effect size d (Cohen, 1992) and a two-sided permutation test (Fisher, 1937) for statistical significance were used to compare the first, second and third recording nights regarding data transmission.

Next, all nights with good transmission quality were further evaluated. First of all, all channels of all recordings were analyzed regarding their data quality: Every channel of every recording was visually inspected, whether any periods of time with bad data quality occurred, and if so, how long they were. The post_analysis.py script was used for this. Next, it was summarized for each of the categories “EMG”, “EOG”, “EEG” and “ECG”, how many channels delivered a good signal in at least 90% of the night. Moreover, the ratio of recordings was calculated for each category, in which at least one channel recorded a good signal in at least 90% of the night (both for only the nights with good signal transmission, and in total over all second and third nights). Finally, the categories were compared between the nights (number of channels with good signal in at least 90% of the night) using Cohen’s effect size d and two-sided permutation testing for significance.

Sleep quality and comfort of the system

The answers of the questionnaires (Schlaf-Fragebogen A and the self developed questionnaire about the system) were evaluated regarding differences in the sleep quality (SQ), the feeling of recovery in the morning (FoR) and the comfort of the system between all three nights, as well as compared to reference groups provided by the questionnaires (SQ and FoR). The sleep quality and feeling of recovery were evaluated as described in the manual to the Schlaf-Fragebogen A (sleep quality items: time to fall asleep in the evening, the number of awakenings during the night, the length of the awake periods during the night, and a self-assessment of the subject, how constant, deep, restless, relaxed, undisturbed and good the sleep of the previous night was, into account; feeling of recovery items: ampleness of sleep and whether, subjectively, the subject felt even-tempered, drowsy, cheerful, fresh, relaxed, and well-rested the next morning). The comfort of the system was based on the single questions of the self developed questionnaire. Descriptive and inferential statistics applying Cohen’s effect size d and two-sided permutations tests for significance were used.

Automatic sleep scoring algorithm and adjustment to the individual subject

Using the plots generated by the script plot_night_predictions.py (for details about the plotting script, please see chapter *3.1.2.2.4 Data visualization and data analyses*), it was first analyzed, whether the system was online and predicted sleep stages during every second of the recorded nights. Next, it was evaluated using the plots of the aforementioned script and the whole-night sleep EEG time-frequency plot of the post_analysis.py script, in how many nights the uncorrected sleep stage probabilities led to a plausible hypnogram. The same analysis was conducted for the sleep stage probabilities, which were adjusted based on the subject-individual sleep characteristics in real-time during the night, and for the post-optimized automatic sleep staging. Furthermore, it was evaluated separately, whether an automatic adjustment to the individual subject was visible at all, and whether it was of benefit, by visually inspecting the generated plots of the nightly predictions.

N3 and REM sleep differences

Using the answers of the nightly questionnaire regarding the dream content (“What describes your thoughts well, which you had before you woke up?”: “rather static thoughts, no story”, “lively or with an action-packed plot”, “it was bizarre”, “it was entertaining; I’d like to continue dreaming about it”. All four items were rated on a scale from “0-not at all” to “7-very much”.), differences between the two sleep stages REM and N3 were determined by calculating Cohen’s effect size d and checking for statistical significance using a two-sided permutation test. The sleep stage was determined by (human) visual scoring using the plotting features of the post_visualization.py script.

Furthermore, the same statistical procedure was used to find out, whether subjects woke up faster from N3 or from REM sleep. The answers to the question “Which letter did you hear first after waking up?”[17] were used for determining the time needed to wake up. It was looked up for every nightly wake-up procedure, which number the letter had, which was written down by the subject.

17 Details about the idea to wake up the subjects using a sequence of acoustic letter stimuli are described in chapter *3.1.2.2.1 Autonomous, easy to program experiments* (*Table 3.2*).

Sleep depth

Furthermore, for the first, second and third night it was analyzed, how quickly the subjects woke up by the nightly stimuli, with the hypothesis that during the second and third night the subjects slept deeper and woke up later, because they became used to the sleep experiment situation. Pearson's correlation coefficient was calculated for the night number (1-3) and the number of the first heard letter (1-6). A one-sided permutation test was applied to test for significance.

The same type of analysis was carried out to investigate, whether the subjects woke up faster during later parts of the night (i. e. more towards the morning).

Workload of the experimenter and costs of the experiment

The actual tasks, which remained to be done by the experimenter and were not carried out by the Traumschreiber system, were briefly reported. Moreover, the workload of the experimenter and the costs of the actual experiment were summarized.

3 Results

3.1 Development

3.1.1 Overview

After about 3 years of development, the requirements described in chapter *2.2.1 Defining the requirements towards the Traumschreiber system* are fulfilled. The Traumschreiber system contains a battery-driven hightech sleep mask, which can be connected via bluetooth low energy to an experiment station – a Raspberry Pi 3 Model B minicomputer. The Traumschreiber sleep mask measures the electrophysiological sleep data of the subject on eight channels using single use sticky electrodes. The data is stored and processed in real-time on the minicomputer using Python scripts – including the possibility to use advanced AI algorithms like deep neural networks based on Keras and Tensorflow. During experiments, the Traumschreiber experiment station can instruct and stimulate the subject via two small USB speakers. A simply adjustable xml script can be used to program complex sleep experiments into the system in a very short time. An advanced data visualization and analysis tool as well as several other python scripts make data standardized analyses quick and easy.

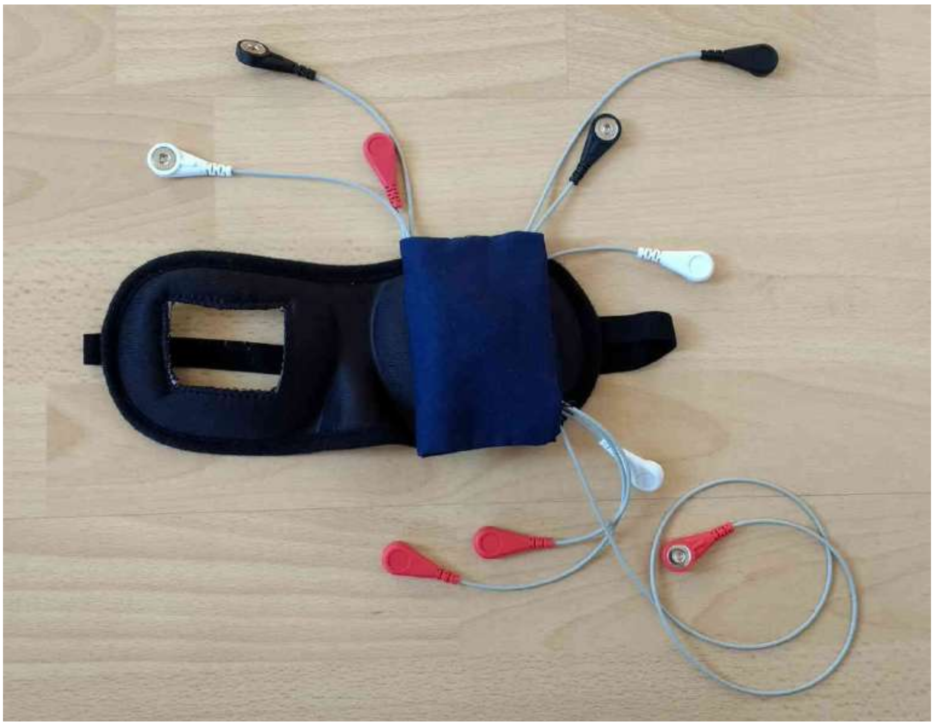


Figure 3.1: The final version of the Traumschreiber sleep mask.

Figure 3.2: One complete Traumschreiber system experiment box, including the sleep mask, the Raspberry Pi minicomputer, the speakers, electrodes and questionnaires. Each subject receives one box, records polysomnographic sleep data at home following the instructions of the Traumschreiber system, and returns the box including the data to the experimenter. No experimenter is needed at the place of recording.

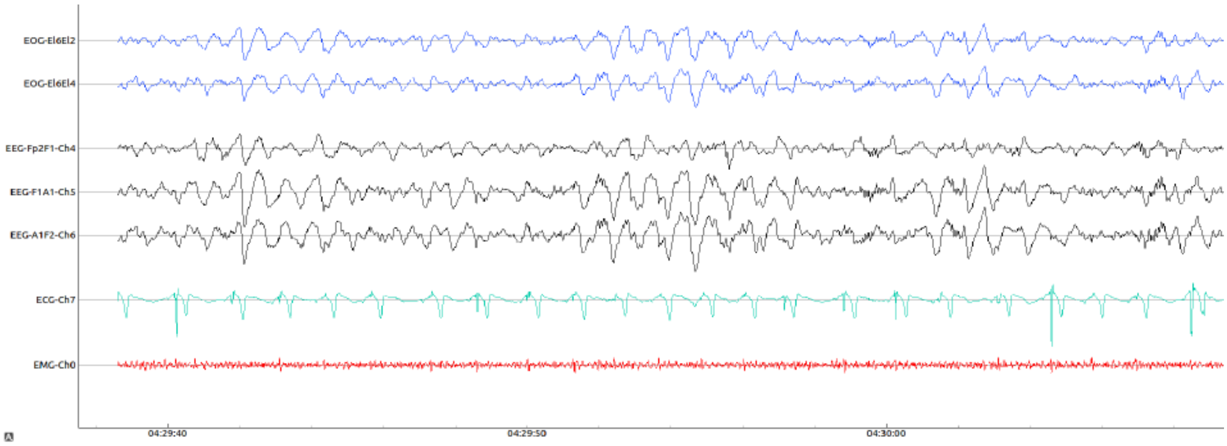


Figure 3.3: N3 sleep (slow wave sleep, SWS) recorded by a subject at home with the Traumschreiber system. Note the clearly visible slow waves.

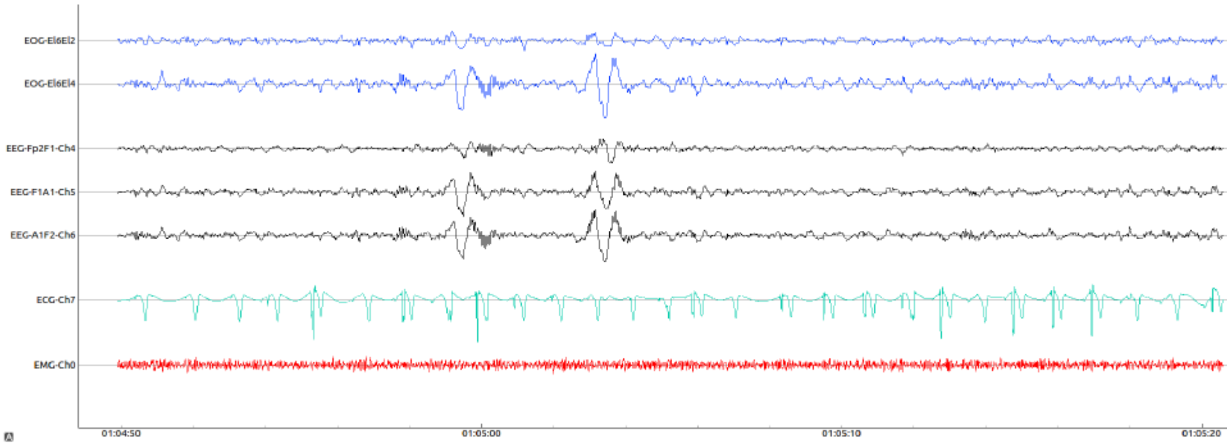


Figure 3.4: N2 sleep recorded by a subject at home with the Traumschreiber system. Note the clearly visible sleep spindles and K-complexes.

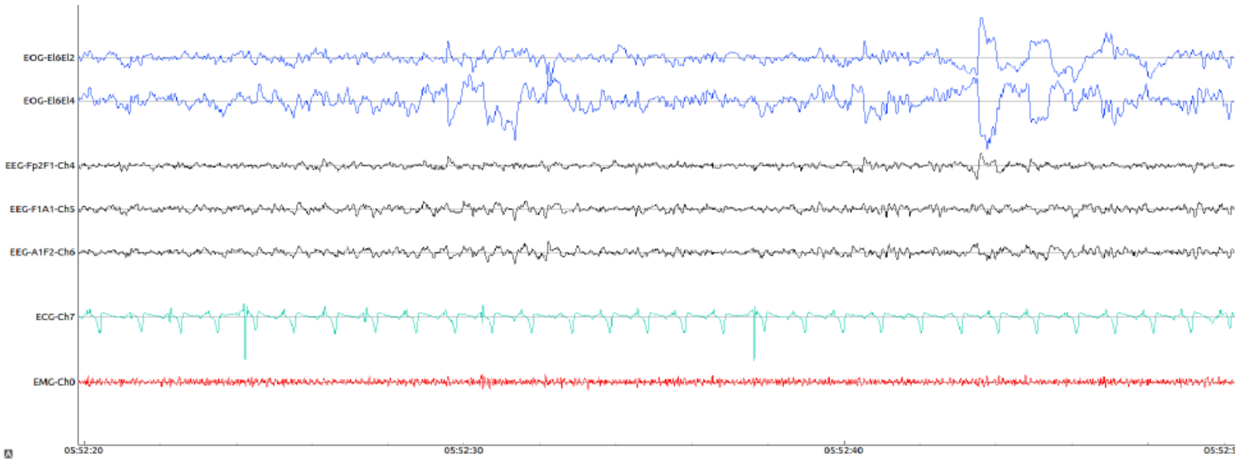


Figure 3.5: REM sleep recorded by a subject at home with the Traumschreiber system. Note the clearly visible rapid eye movements.

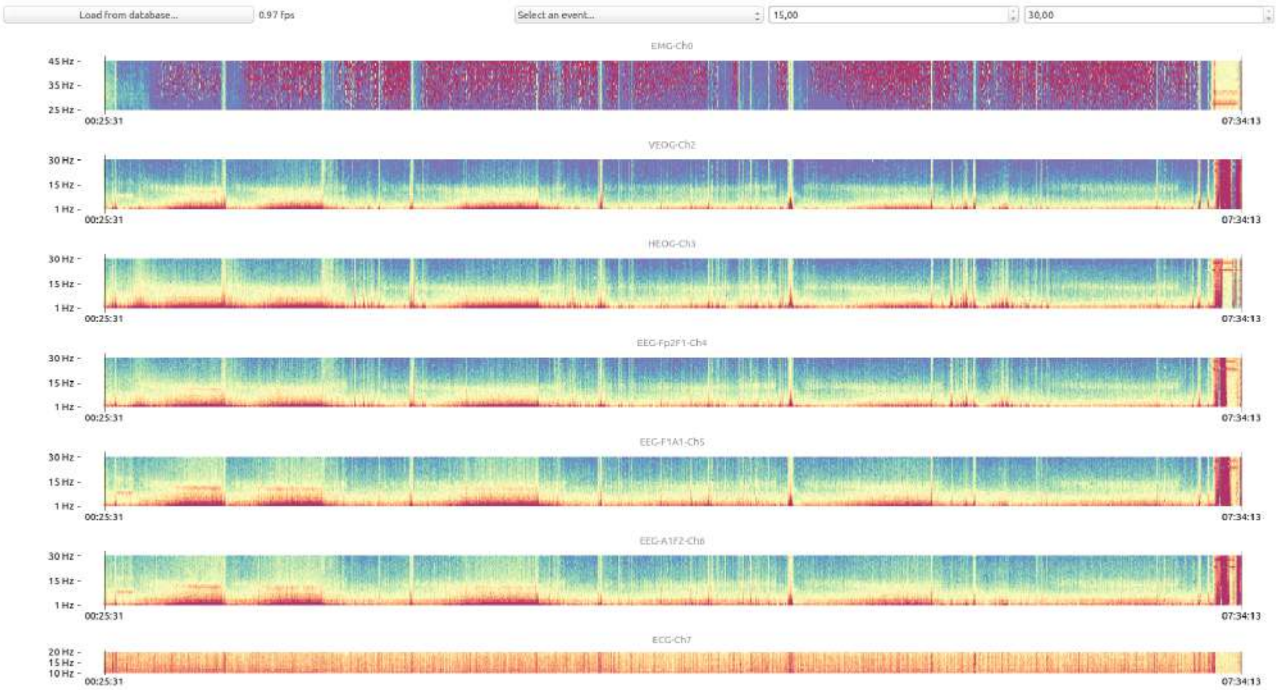


Figure 3.6: Time frequency plots of a whole night recording of the Traumschreiber system of all relevant channels (the channel between the chin and the bottom eye electrode is not depicted). Note the clearly visible activity in the slow wave and the sleep spindle frequencies, and the arousal/movement artifacts, which give the sleep researcher an instantaneous overview of the night.

3.1.2 The components of the Traumschreiber system

3.1.2.1 Hardware

3.1.2.1.1 Sleep mask

The sleep mask is worn on the eyes and is based on a black cotton-synthetic sleep mask with an adjustable, elastic band. A 5x4 cm hole is cut into the right eye of the sleep mask, enabling the subject to look through the sleep mask. A 10x7cm red or blue piece of cloth (a pocket) is sewn on the front of the left eye, which contains and protects the electric parts of the sleep mask, i. e. the PCB. This cotton bag can be opened and closed with a hook and loop fastener, and is connected to the sleep mask with four eyelets. The board lies inside, and is placed in a way that the two LEDs, which are mounted on the PCB, can shine through two of the four eyelets, ready to stimulate the (closed) eyes during the night.

Ten single use sticky electrodes can be connected to the board via cables of different length. In the suggested configuration, two cables are connected to electrodes on the subject's chin (measuring EMG), three cables are connected to electrodes outside and below the left eye, outside and above the left eye, and outside and above the right eye (measuring VEOG and HEOG), three cables are connected to two electrodes on the left and right forehead as close to the hairline as possible and to one electrode on the mastoid behind the left ear (measuring EEG), one cable is connected to an electrode on the left side of the chest (measuring ECG), and the last cable is connected to another electrode on the forehead, which is used as a ground electrode. The cables are adjusted in length and soldered onto the PCB so that they can easily reach the electrode locations, without being too long – thus, reducing the weight of the sleep masks and increasing comfort.

The cable design and the electrode locations can be arbitrarily changed by using different cable lengths and sticking the electrodes to other hairless parts of the skin, allowing for different recording locations, if needed. Table 3.1 displays the channels which are recorded in the suggested configuration. This configuration is suggested as it records EMG, EOG and EEG, which is needed for standard polysomnographic recordings, and additional ECG, which is useful for example for heart rate variability analyses. The channels can be combined, i. e. it is possible to re-reference the electrodes to one common reference.

Electrodes	**Positions**	**Channel number**	**Suggested use**
0 and 1	Chin, chin	0	EMG
1 and 2	Chin, below left eye	1	-
2 and 3	Below left eye, above left eye	2	VEOG
3 and 4	Above left eye, above right eye	3	HEOG (+EEG)
4 and 5	Above right eye, left forehead close to hairline	4	EEG (Fp2-F1)
5 and 6	Left forehead close to hairline, left mastoid (behind the ear)	5	EEG (F1-A1)
6 and 7	Left mastoid (behind the ear), right forehead close to hairline	6	EEG (A1-F2)
7 and 8	Right forehead close to hairline, left side of the chest	7	ECG

Table 3.1: Overview of the electrode positions and recorded channels in the suggested recording layout.

The PCB is equipped with a holder for a 3.6V lithium-ion rechargeable coin cell battery, and a micro-USB recharging plug. The sleep mask can thus be recharged using a standard smartphone recharging cable, or the power cable of a Raspberry Pi.

Furthermore, the PCB has a power switch, making sure that no battery power is used, if the sleep mask is not in use. Both the recharging plug and the power switch are hidden inside the cover bag, but can be reached with a little bit of fiddling without opening the cover bag.

Furthermore, the PCB is equipped with dozens of electric components, ranging from simple resistors to microcontrollers. Using the data flow from the electrodes to the minicomputer as a metaphor, the differential signal between two electrodes is first amplified by a factor of 100 by instrumentation amplifiers, then passes hardware-based filters (an active 2nd order Sallen-Key high-pass filter with a gain by a factor of 10 and a cutoff frequency of 0.05 Hz, and a passive 1st order resistor-capacitor low-pass filter with a cut-off frequency of 72 Hz, resulting in an attenuation of the signal by a factor of 10 at roughly 720 Hz (plots of the filter response in the appendix)), is then again amplified (by a factor of 1, which is adjustable in software) and sampled by an analog to digital converter (ADC) at 244 Hz with 12 bit resolution. Both the analog to digital conversion and the last amplification are done by a microcontroller (Atmel XMEGA8E5). The obtained values for each channel can be encrypted by a dedicated chip in hardware (model ATECC508A, not active at the moment) and are then passed via UART to a bluetooth chip (Broadcom BCM20737S), which transmits them via bluetooth low energy to arbitrary BLE-able devices, e. g. smartphones or, as used in the final version of this system, to a Raspberry Pi 3 Model B minicomputer. In case the BLE signal is lost, for example if a subjects goes to the toilet, the connection is automatically re-established once the signal is in reach again. The BLE chip can also receive commands from the paired device, e. g. for setting the color of the two LEDs to different colors. Several pins on the microcontroller are not in use so far, making it possible to add further customized features, e. g. connecting and controlling other stimulating devices. Further technical specifications can be found in the online appendix.

3.1.2.1.2 Experiment station

Besides the sleep mask, a Raspberry Pi 3 Model B minicomputer with the linux derivate Raspbian Noobs as operating system is the second essential part of the Traumschreiber system. The Raspberry Pi is connected to two small USB speakers (Trust Leto 2.0 USB), for which it uses an external USB sound card in order to improve sound quality, since the standard audio jack of the Raspberry Pi produces a noisy audio signal. The minicomputer guides the subject through the experiment by giving acoustic instructions of what to do when. Moreover, the minicomputer saves and analyzes the incoming data from the sleep mask in real-time.

In principle, a smartphone could be used as an experiment station, too. This option was investigated in a bachelor thesis project of Martin Jäkel (Jäkel, 2017, supervised by the author of this dissertation). However, it became apparent, that developing a complex system like the Traumschreiber system for a smartphone is much more complicated and takes much longer time than for a linux-based minicomputer. As a result, despite the promising impressions of the developed Android app prototype (see figure 3.7 for screenshots, complete code in the online appendix), this option was postponed to future developments.

Following this line of thought, the direct data transfer of the recorded data via smartphone over the internet to a web server was briefly investigated as well. The web server analyzes the data in real-time and streams back its real-time calculations to another device – e. g. the sleep researcher's computer (web browser). The idea behind would be to enable the sleep researcher to see the recorded data of all subjects in real-time, or possibly to even combine the data of multiple subjects in new experimental paradigms. Even though this direct data transfer via internet worked well in a demo case, this topic was not investigated further.

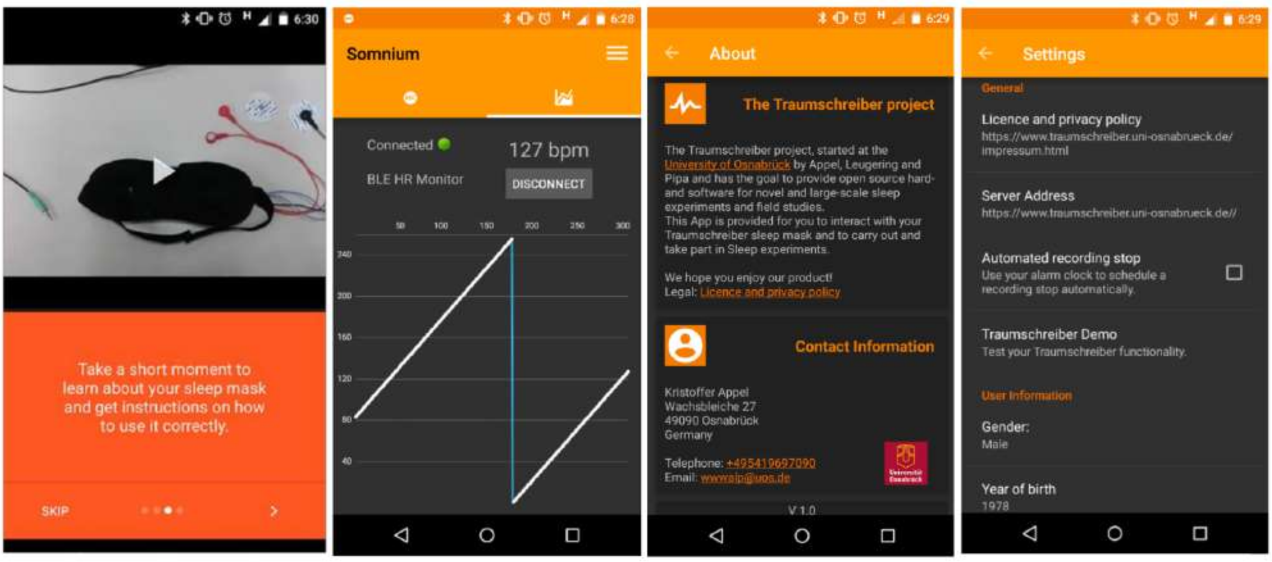


Figure 3.7: Four screenshots of the prototype Android app, demonstrating that an Android smartphone could be used as the experiment station instead of the Raspberry Pi minicomputer, if this line of development is further pursued (image source: (Jäkel, 2017)).

3.1.2.2 Software

3.1.2.2.1 Autonomous, easy to program experiments

Several python scripts are stored on the minicomputer, some of which are executed during startup. The user (experiment subject) only needs to connect the minicomputer to a power plug. After about 30 seconds, the Traumschreiber software is started and the user is welcomed via the two speakers. What happens next, can be specified in the experiment xml protocol file (see figure 3.8 for a simple example, and table 3.2 for further explanation). In a basic experiment setup, the user first receives some acoustic instructions (e. g. to put on the sleep mask and to switch it on), before going to bed and conducting the sleep part of the study. Every experiment action, which the Traumschreiber system carries out according to the experiment xml protocol after the sleep mask is switched on, is time logged, and can be looked up later during data analysis.

```
<root>
    <experiment>

        <sound>welcome.wav</sound>
        <sound>follow_evening_instructions_then_turn_on_TS.wav</sound>
        <wait_for>TS_connected</wait_for>
        <sound>connected.wav</sound>
        <sound>calibration.wav</sound>
        <sound>good_night.wav</sound>

        <wait_for>sleep</wait_for>
        <wait_for>duration_07:30:00:000</wait_for>
        <sound>alarm_clock_song.wav</sound>

    </experiment>
</root>
```

Figure 3.8: Example xml experiment file. After a welcome message of the Traumschreiber system after startup, the subject is asked to follow the instructions from the video tutorial and the subject information leaflet. In these instructions, the subject is told how to put on the sleep mask and when to switch it on (when being ready for sleep). Next, when a connection to the sleep mask was established, the subject is informed about this and is asked to conduct the bio signal calibration. Afterwards, the subject is asked to sleep. Next, the Traumschreiber system waits for sleep onset, i. e. that the internal real-time sleep classifier detects either REM, N1, N2 or N3 sleep. 7.5 hours later, the Traumschreiber system wakes up the subject with a melody.

xml element	Function
<root>, </root>, <experiment>, </experiment>	Setting up the experiment xml file.
<sound>abc.wav</sound>	Plays the sound file “abc.wav”, which has to be located in the experiment folder, via the USB speakers.
<wait_for>abc</wait_for>	Pauses the experiment execution, until the condition “abc” is met. “abc” can have various values: • “duration_12:34:56:789”: waits for an arbitrary period of time, in this example 12 hours, 34 minutes, 56.789 seconds • “TS_connected”: waits with any further experiment execution until the sleep mask is connected • “sleep”, “N1”, “N2”, “N3”, “REM”, “wake”: waits until the Traumschreiber system has detected sleep or the sleep stage specified • “LRLRLR”: waits until the Traumschreiber system has detected a left-right-left-right-left-right eye movement pattern It is possible to combine these elements using “OR”. For example, if “abc” is set to “N3 OR REM OR duration_00:30:00:000”, the Traumschreiber system waits until either N3 sleep or REM sleep have been detected or 30 minutes have passed (whatever happens first). Only then the next xml element is carried out.
Customized elements, e. g. <arousal_procedure> or <coding_number>	Any customized element can be used for advanced experiment control, if the action is specified in the python file “eese.py” in the function “action”. For example, the xml element <arousal_procedure> evokes one such customized function. In this function, every five seconds one of the letters 'o', 'b', 'r', 'i', 'h', 'j', 'l', 'v', 'm', 'ä' is randomly selected as an acoustic stimuli, with increasing volume, in total five times. This is followed by six alarm clock-like beeps. This procedure was used to wake up the subjects and to let them write down, at which letter they woke up. As a second example, the xml element <coding_number> leads to the playback of the next element in a list of acoustic stimuli ('4', '9', '6', '7', '8' and '5'). i. e. every time the xml element is reached during the experiment, another number is presented to the subject. This was used to check, whether a subject really woke up and filled out the nightly questionnaire, or just invented some answers the next morning.

Table 3.2: Detailed description of the experiment xml file commands.

3.1.2.2.2 Automatic sleep staging and pattern recognition

The Traumschreiber system supports a simple form of real-time sleep stage classification, and can adjust the experiment accordingly. For example, it is possible to wait for the occurrence of REM sleep and then playback a sound in order to wake up the subject (e. g. for REM sleep deprivation studies, or for collecting REM dream reports). For this, Keras (Chollet, 2015) with Tensorflow (Abadi, Barham, et al., 2016) backend is used, which is one of the best and most advanced frameworks for machine learning using deep neural networks (Kovalev, Kalinovsky, & Kovalev, 2016)[18]. The neural network classifies each second of sleep recording as either wake, N1, N2, N3 or REM sleep. The Traumschreiber system then averages the classifications over the last 20 minutes, and calculates which sleep stage was most prominent in this time window. Based on the individual sleep characteristics of the subject in previously recorded nights and a calculation, how much the distribution of classifications of each night differed from the average healthy human sleep stage distribution found in literature (Rama & Zachariah, 2013), a correction term is then applied to the data. This enables the Traumschreiber system to adjust the sleep stage classifications subject-individually in real-time. For example, if less N2 sleep was detected by the Traumschreiber system for a specific subject in previous nights than would be expected for healthy human subjects, an additive bias term is applied to make the live sleep stage classification more likely (assuming the subject has a healthy sleep). In this example, it might be the case that for this very subject, less EEG activity in a specific frequency band is measured, making the neural network select the wrong sleep stage more often. For a critical review, of how well this procedure works in practice, see chapter *3.2.4.6 Automatic sleep scoring algorithm and subject-specific adjustment*.

A similar simplistic neural network-based approach as for classifying sleep stages in real-time is implemented into the Traumschreiber system for detecting specific graphoelements during recording, e. g. specific eye movement patterns like the lucid dreaming eye signal left-right-left-right-left-right[19]. Note, however, that this real-time pattern recognition algorithm was trained only on few training examples of one subject (the author of this dissertation) and needs further improvements, before it can be used in autonomous experiments. See figure 3.9 for an example of the real-time eye movement detection.

18 Further details on the technical implementation can be found in the appendix.

19 Further details on the technical implementation can be found in the appendix.

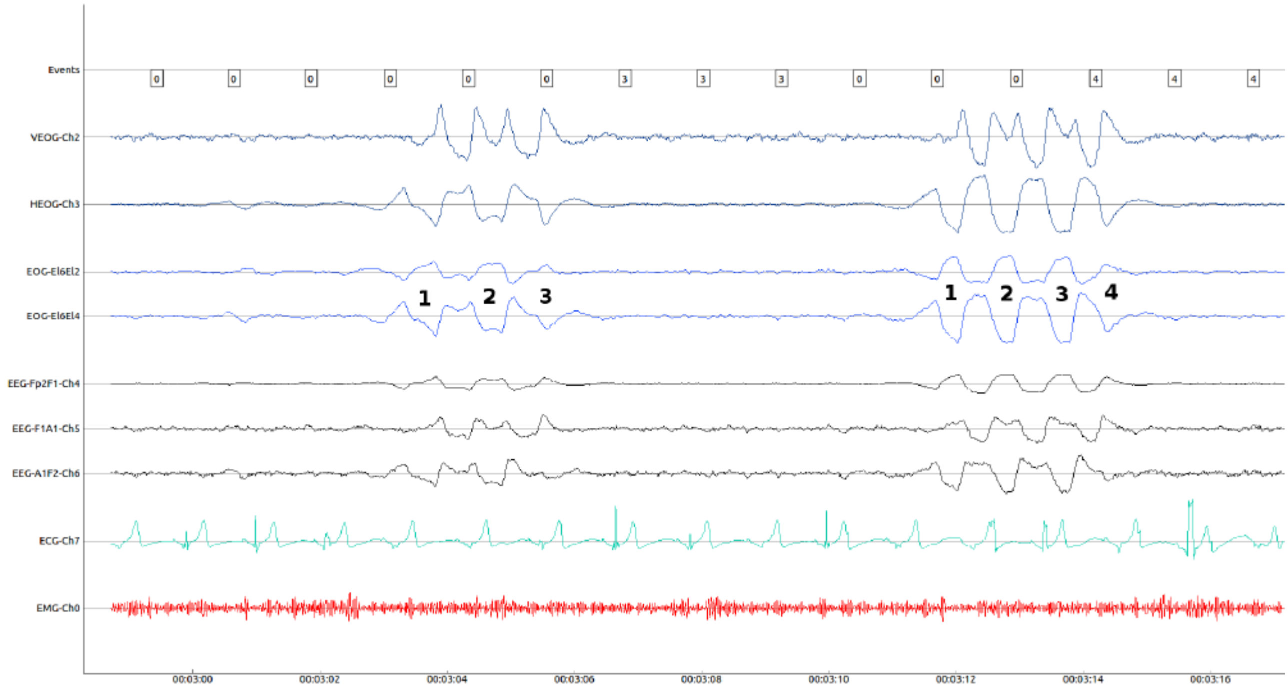


Figure 3.9: Example of the real-time pattern detection using a simplistic Keras/tensorflow neural network approach. Note the 3xLR eye movement of the subject at 00:03:04 and the corresponding real-time classification "3" after the eye signal (top row), and the 4xLR eye movement at 00:03:12 and the corresponding real-time classification "4", as well as the "0" classifications for the rest of the time.

Using state-of-the art machine learning technology, the Traumschreiber system is thus able to not only analyze the macro structure of sleep in real-time (i. e. sleep stages) – it can also detect patterns in the data in real-time. It has to be emphasized though, that it was not the goal (yet) to optimize the performance of the machine learning algorithms, not only as not enough recorded nights are available as training data at the moment. The neural network within the Traumschreiber system should be seen as a placeholder for more elaborated, better trained deep neural networks. Any more sophisticated deep neural networks can then be used instead of the simplistic approaches for sleep staging or pattern recognition developed up to now. They can rely on the same hardware and software layout – future developers only need to exchange two files: the serialized JSON description of the network structure and the serialized weight matrix of the network. In case the input data structure to the network is different, a few more lines of code need to be adjusted in the python file classification_methods.py.

3.1.2.2.3 Real-time data processing using encapsulated, modular scripts

The software is designed that the data coming from the sleep mask via BLE are received by a Python script, which directly broadcasts them using a socket server via the user datagram protocol (UDP). This enables arbitrary many listeners (clients) to receive, store or process the data further – also in independent processes and (in principle) also in any arbitrary programming language that supports sockets and UDP. A second UDP server is created automatically during start of the system, which forwards control messages between all different software modules, enabling a communication between arbitrary scripts written in arbitrary programming languages and running in several processes.

In the final version of the Traumschreiber system, one such modular script (save_data_to_database.py) saves the raw data every 30 seconds into an sqlite database, i. e. the data is stored safely on the microSD card even if the minicomputer is abruptly disconnected from power supply. An experiment subject can thus just unplug the minicomputer in the morning without having to shut down the system.

Another modular script (replay.py) is able to simulate a recording: by sending data via the UDP port, all other software modules can be tested for functionality also during the day, which makes the software development much easier.

A third modular script serves as an event logger. It listens to all messages, which are forwarded by the control server, and identifies events, which are then saved into the sqlite database. That means, that any module can directly save any marker together with a time stamp into the database, without having to know anything about the underlying database filename, structure or type. This feature is for example used for saving the time stamps of acoustic stimuli. Furthermore, it is also possible to save events manually by using the GUI of the script.

3.1.2.2.4 Data visualization and data analyses

Last, but not least, the Traumschreiber system provides several scripts for data visualization and data analysis – both for real-time visualization during a recording, and for later post-analysis.

Data transmission analysis

The stand-alone python script recording_transmission_analyses.py calculates for each recorded database in a given folder,

- when the recording started (timestamp of first data point), and when it stopped (timestamp of last data point),
- the total recording duration, i. e. the difference between the first and last datapoint,
- whether the recording has a total recording duration of less than 5 minutes and should be classified as “mistake”, e. g. for when the subject mistakenly started the experimental recording due to having switched on the sleep mask during preparation and needs to replay the calibration later,
- whether the recording has a total recording duration of more than 5 minutes (classified as “valid recording attempts”),
- the ratio of seconds with no or with bad data transmission (i. e. less than 50 or 232[20] data points, respectively), and the ratio of seconds with good data transmission (i. e. at least 232 data points),
- classified the recording into one of the categories “mistake”, “valid recording attempt with signal loss”, “valid recording attempt with bad signal transmission”, and “valid recording attempt with good signal transmission”, depending on the ratio of no, bad and good seconds regarding the transmission quality,
- automatically finds the related temperature log file for the Raspberry Pi and obtains the temperature development throughout the night, and the time duration the Raspberry Pi was listening for data packages.

The script then plots a data transmission diagram for each database file, in which one can see, how the transmission changed in the course of night. Furthermore, the most important statistics regarding data transmission success and the Raspberry Pi temperature and online time are displayed. Moreover, the script saves the calculated data in a summary csv file, which can then be used for further analyses. Please see figures 3.10 and 3.11 for exemplary plots of a recording with good signal transmission and a recording with signal loss.

20 Why 232? In optimal case, 244 data points are transferred on average every second. Bad data transmission was defined as more than 5% of data loss during data transmission, i. e. less than 232 data points per second.

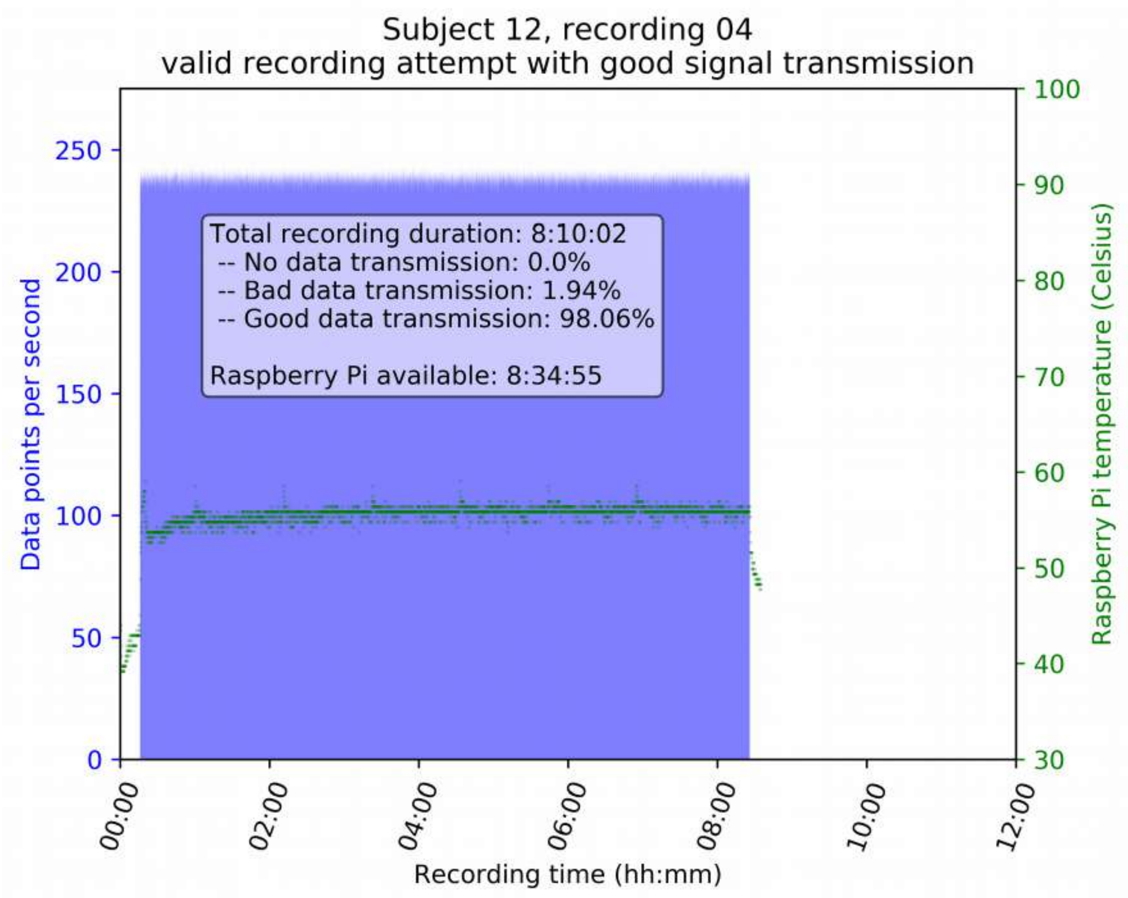


Figure 3.10: This plot, which was generated by the recording_transmission_analyses.py script of the Traumschreiber data analysis package, shows that the signal transmission rate stayed constantly high throughout the whole night, close to the optimum of 244 data points per second. Only for a small amount of time, the signal transmission was bad, i. e. more than 5% of the data points of a second were lost. It can further be seen, that the sleep mask was switched on and started sending data about 15 minutes after the Raspberry Pi was switched on, and switched off a about 10 minutes before the Raspberry Pi. The temperature of the Raspberry Pi stayed constantly at around 57 °C during recording, showing that the Raspberry Pi did not overheat due to potentially too heavy calculations.

Real-time data visualization

The script “standalone_plotting.py” enables the researcher to quickly look at the recorded data in real-time. All eight channels can be bandpass filtered and are displayed (last 10 seconds). This script can be used as a reference of a minimal data processing pipeline, as it also includes the BLE handling and plotting of the raw signal.

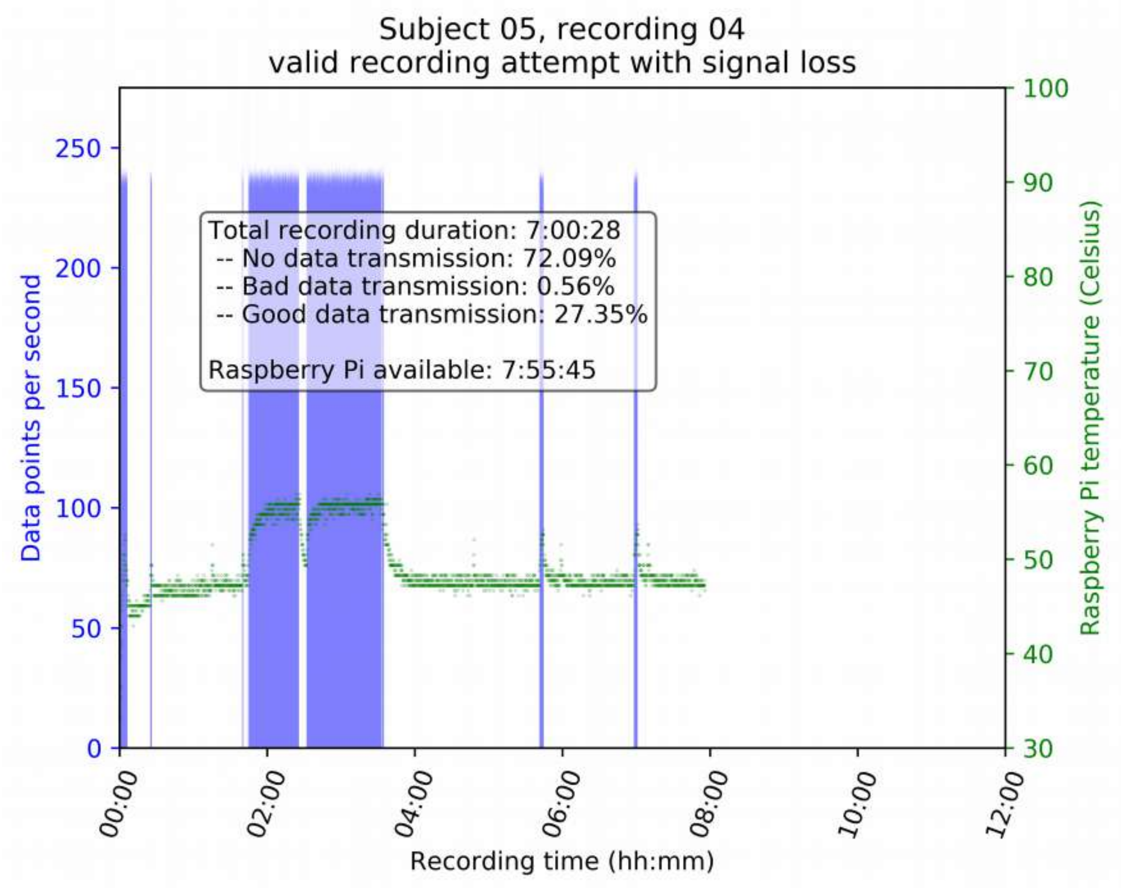


Figure 3.11: This plot, which was also generated by the recording_transmission_analyses.py script of the Traumschreiber data analysis package, shows that no data was transmitted in this recording throughout most parts of the night. The Raspberry Pi was switched on the whole time, as can be seen at the green dots indicating its temperature. Furthermore, it can be seen, that the temperature of the Raspberry Pi went up by 10°C, when it received data packages, but stays far below the threshold of 80°C, at which it reduced the CPU speed in order to prevent overheating. Thus, it seems unlikely, that the data loss of this recording was a reception problem of the experiment station. It is more likely, that the sleep mask did not send data throughout the whole night, probably due to a loose on-off power switch of the sleep mask used in this recording.

Visualizing and manually sleep scoring a recording

A recorded night can be viewed and manually scored by running the script post_visualization.py. After starting the script, a GUI opens and a database can be selected by clicking a button. The data are completely loaded into RAM, which is why this script cannot be executed on a Raspberry Pi 3 Model B (too few RAM for a whole night recording). The advantage of loading the whole data into RAM is that afterwards one can jump to any point in the recording without waiting time.

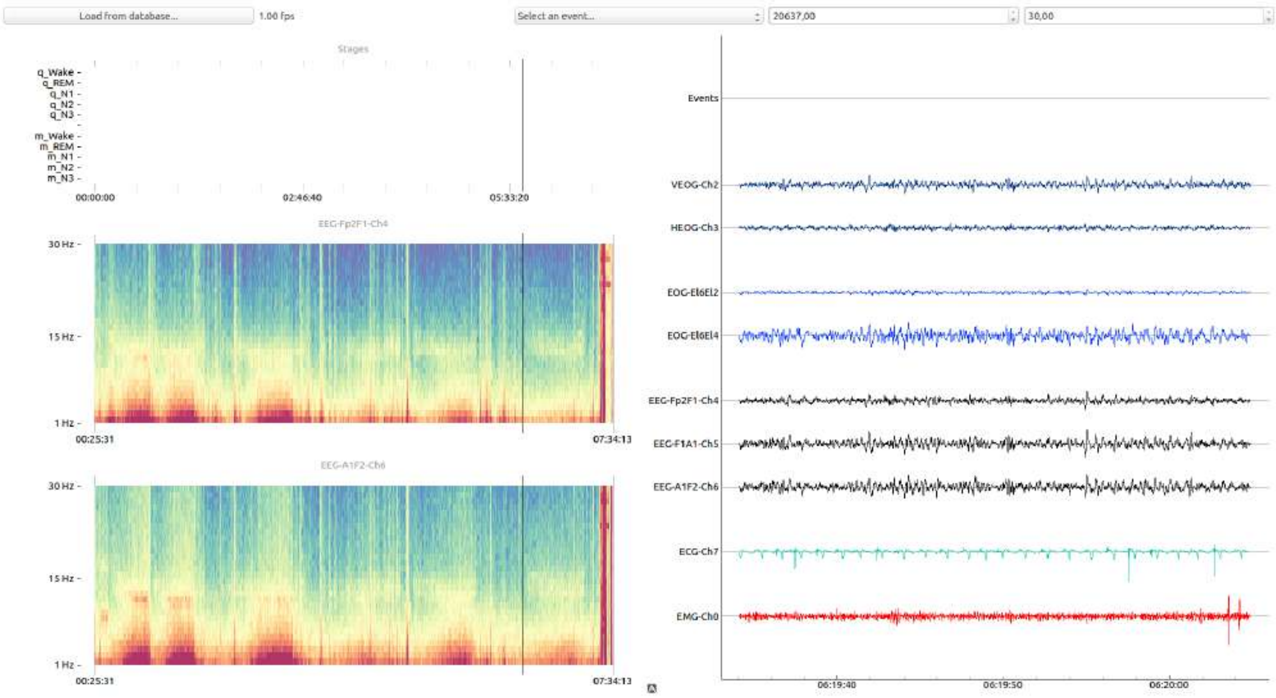


Figure 3.12: The GUI of the post_visualization.py script after loading the data, which was recorded autonomously at the subject's home by the Traumschreiber system. The left side shows an overview of the complete night, here two EEG channels are displayed in a time frequency plot. Above is the (still empty) sleep staging plot. The right side depicts a detail view of 30 seconds length (length can be adjusted in the top right corner). The black lines in the left plots indicate, which time point is selected for the detailed view.

On the right side of the GUI, the raw data are displayed, usually in a 30 second window. This is the detailed view of the data, in which the researcher can look for sleep graphoelements like sleep spindles or REMs. Moreover, the events, which took place during this time window, as displayed in the right plot as well. The length of the time window can be adjusted in the top right corner. On the left side, transformed versions of the data are shown, e. g. a time frequency plot using fast fourier transform (FFT) or complex morlet wavelets. This is useful for getting an overview of the night and channels, and with a bit of practice one can identify periods of time with a lot of slow wave activity (N3 sleep), spindle actitivy (N2) or mixed frequency EEG with REMs (REM sleep). Moreover, one can identify arousals, as they cause artifacts in the signal and its frequency transform. When pressing the keyboard shortcuts Y, X, and C, only the left, both or only the right plots are shown, enabling the scientist to inspect the plots in greater detail.

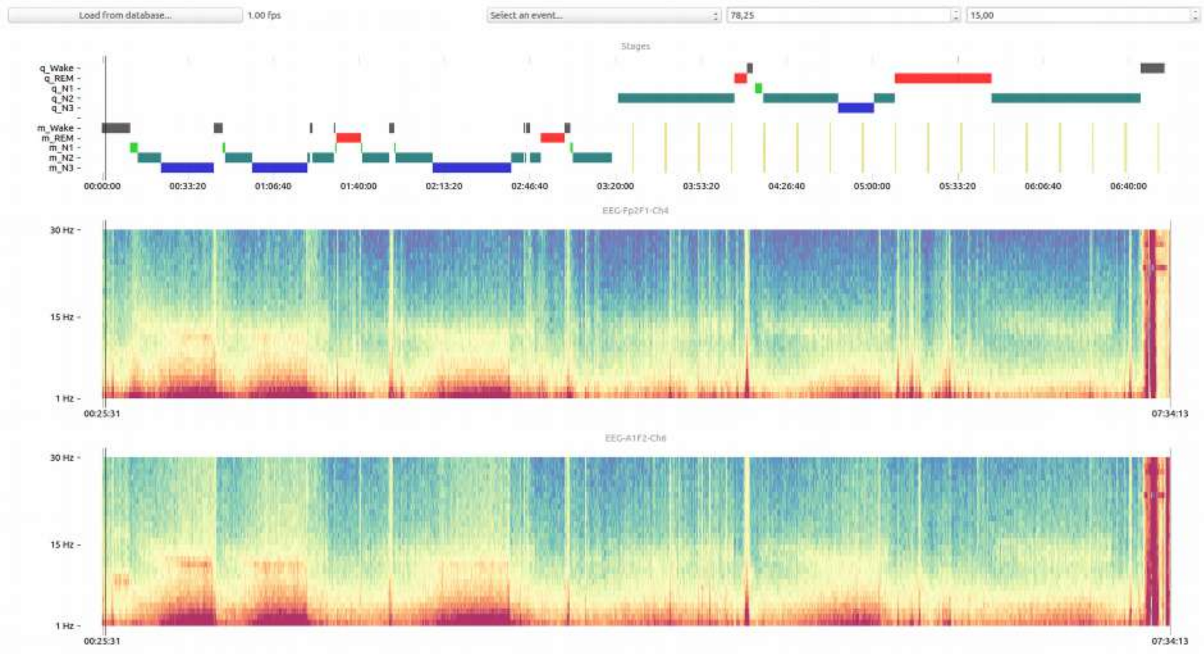


Figure 3.13: After pressing the Y shortcut on the keyboard, the night overview is displayed in full window width. This screenshot was taken after sleep staging (top plot filled, see next pages).

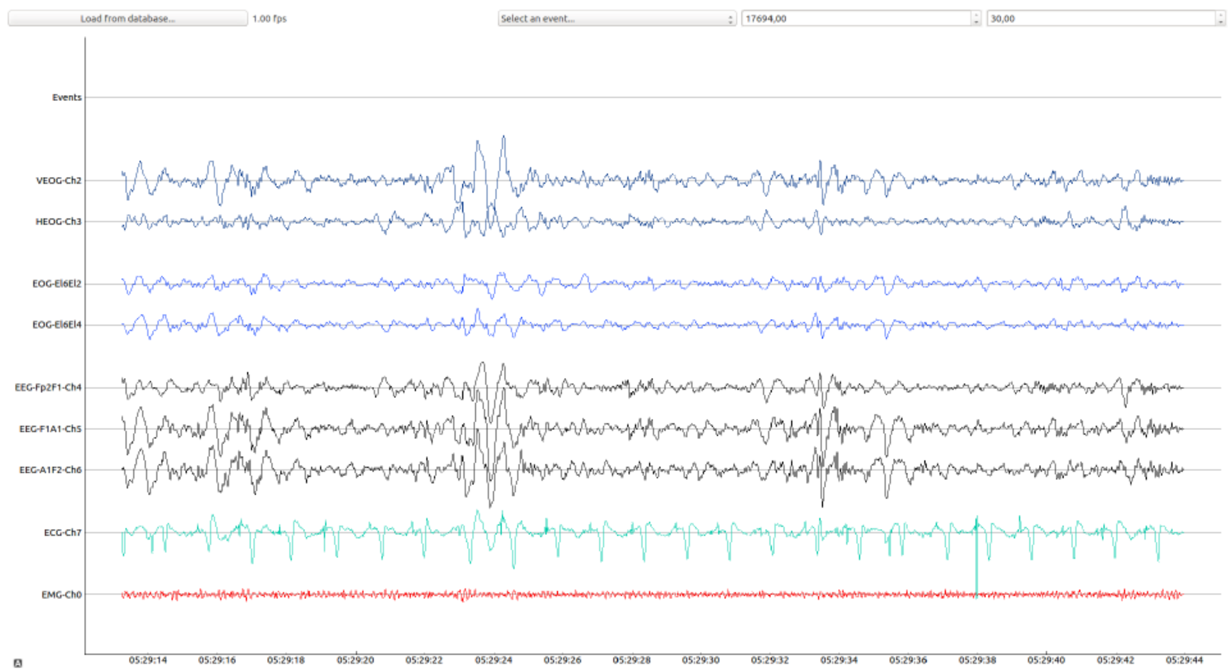


Figure 3.14: After pressing the C shortcut on the keyboard, only the detailed view is displayed, which makes it easier to identify sleep spindles or other sleep grapholements. The following recording channels were selected to be displayed in the config xml file: VEOG and HEOG, a re-referenced channel depicting the voltage between the left mastoid electrode (electrode 6) and the electrode below the left eye (electrode 2), another re-referenced channel (left mastoid vs. electrode above the right eye), three EEG channels, ECG and EMG.

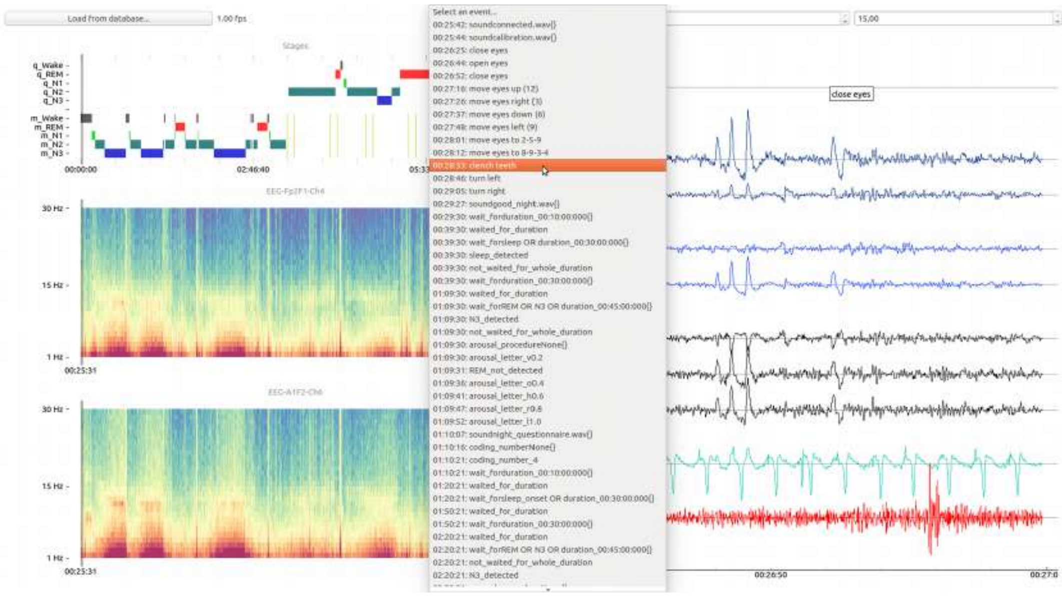


Figure 3.15: The researcher can comfortably jump directly to the time point, at which the Traumschreiber system carried out a specific action or at which other events happened, by using the events dropdown menu.

It is possible to navigate through the data easily in multiple ways. First, the researcher can click into the frequency plots with the left mouse button. The right plots are updated and centered around the selected second from the left plots. Second, one can directly enter either an exact second of the recording or a POSIX timestamp (as used in the database for storing the data and events) in the GUI (top row, second field from the right). Third, one can select one of the events, which have been stored in the database, by clicking on it in the dropdown menu in the top row. The right view is then centered around the timestamp of this event. Last, one can use the keyboard shortcuts N and M for quickly jumping one time window back or next.

The scientist can choose, which channels should be displayed, by modifying a configuration xml file. This also includes combinations of channels, in case one wants to use a common reference for several channels[21]. Furthermore, it is possible to adjust bandpass and median filters for the left and right plots, and to specify plot properties like scaling factor, line color or labels. Lastly, one can also select, which events should be displayed.

21 When combining several channels, be aware that the channels 1, 3, 5 and 7 are inverted due to the technical layout of the instrumentation amplifiers.

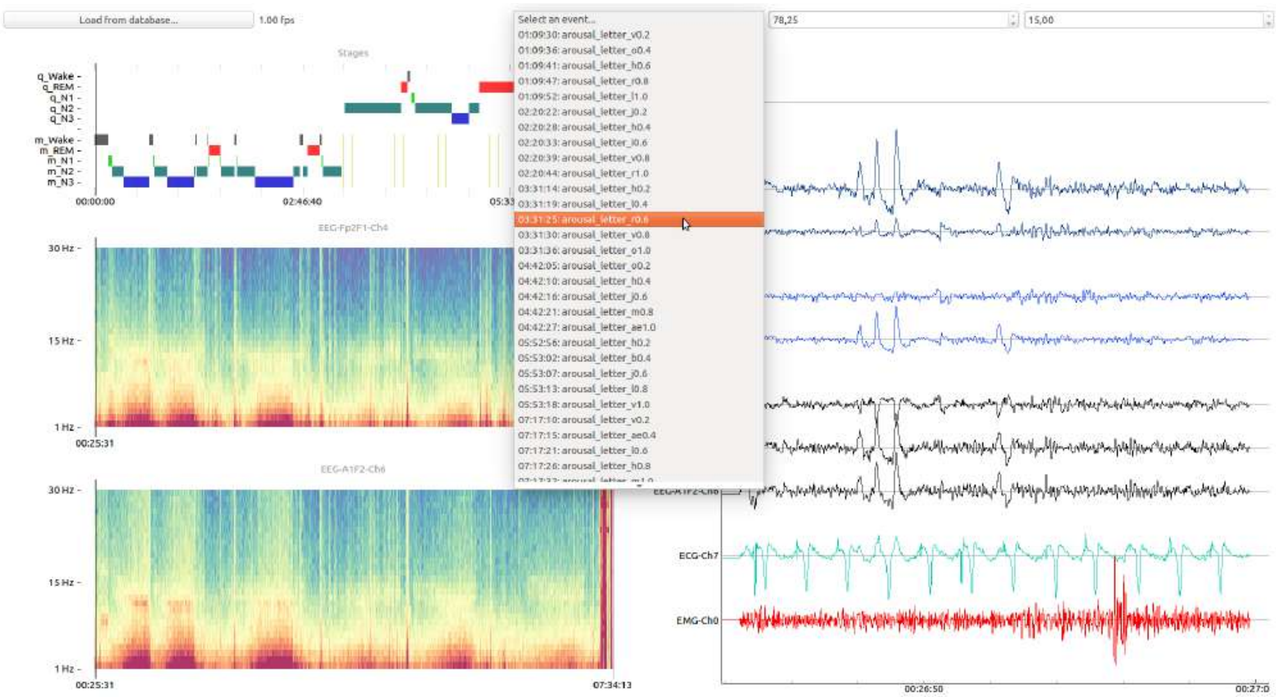


Figure 3.16: When specifying a filter in the config xml file, only certain events are displayed in the dropdown menu and in the detailed view (here: events containing the word “letter”).

The last important feature of the post visualization GUI is, that the sleep scientist can manually sleep stage the recorded data in two different ways. First, it is possible to score the sleep data on a 30 second bases, similar to the procedure in commercial sleep analysis software. The keyboard shortcut A activates the sleep staging. Next, by pressing the keyboard shortcuts W, 1, 2, 3 or R, one can score the currently displayed time window as awake, N1, N2, N3 or REM sleep period, and the focus moves 30 seconds forward, ready for the next scoring. The hypnogram is displayed in the left side plots. It is possible to sleep stage a whole night manually using this standard procedure. However, this time consuming method is not suited for larger datasets consisting of dozens or even hundreds of recorded nights (as might be the case when collecting sleep data with several Traumschreiber systems simultaneously). As a consequence, a second scoring approach has been developed in the Traumschreiber system, named *quick scoring*. By pressing the keyboard shortcut Q, this scoring mode is activated. Next, the sleep scorer presses one of the keyboard shortcuts W, 1, 2, 3 or R, to select the sleep stage to be scored. By pressing and dragging the right mouse button in one of the left plots, the whole selected time period is assigned the according sleep stage (based on one second bins). This way, longer periods of time containing only one sleep stage can be quickly scored at once, e. g. longer periods of slow wave sleep. Questionable time windows can be easily clarified by left-clicking into them and inspecting them the standard way. Especially for creating large training sets for machine learning based sleep staging algorithms, this quick scoring was developed. By pressing the keyboard shortcut S, the manually or quickly scored data are saved into the database.

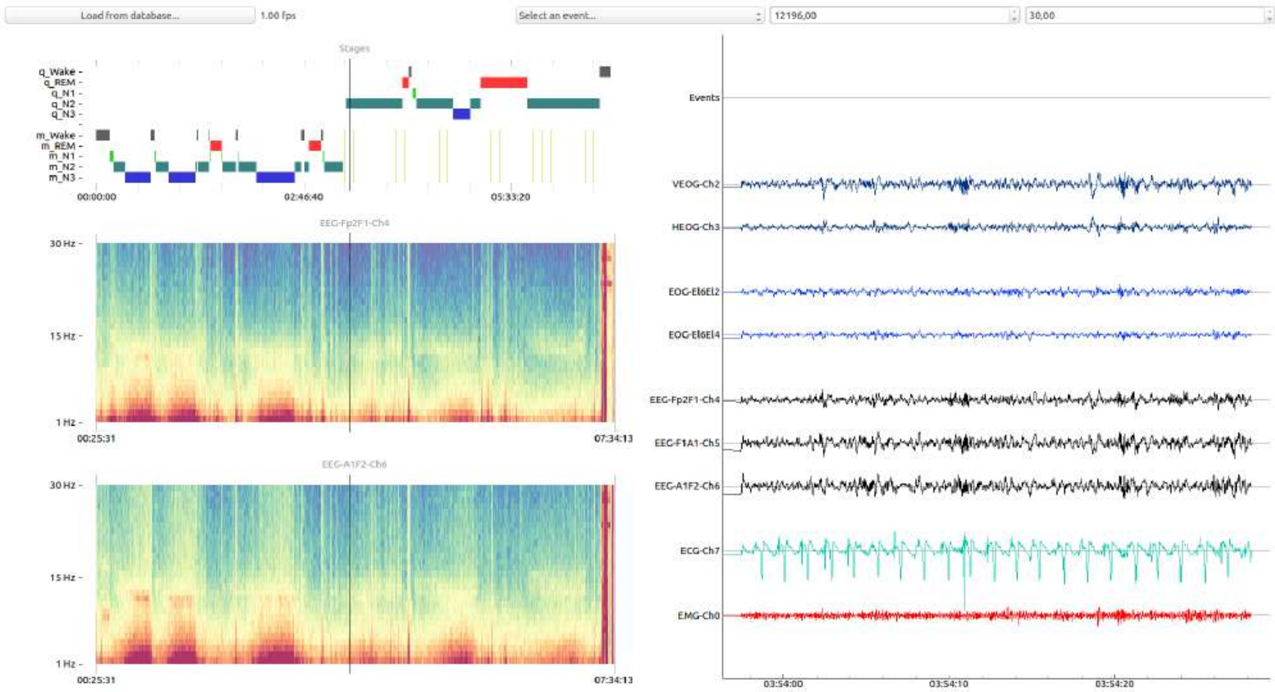


Figure 3.17: The top left plot shows, that the sleep researcher scored the first half of the night in the conventional way using 30-second epochs (sleep stages starting with m_), and the second half of the night with the quick scoring method (sleep stages starting with q_). Both classifications are saved into the database, which contains also the rest of the recording, making them available for training machine learning algorithms such as deep neural networks.

<u>Visualizing the real-time sleep scorings</u>

As described above, the Traumschreiber system is capable of performing real-time classification of the recorded data using a neural network approach based on Keras/tensorflow. The predictions of the AI can be visualized using the script plot_night_predictions.py. The script analyses every database in a given folder, and saves several figures as *.png files.

The first figure plots each prediction timepoint against its POSIX timestamp. There should be one prediction every second, resulting in an (unspectacular) straight line in the plot. If the line is not straight, however, this means that the classification algorithm did not predict a sleep stage every second, i. e. the Raspberry Pi was not able to predict the sleep stage within one second. This was the case at some points during development, and was solved by refactoring the code and minimizing the RAM requirements of sleep stage prediction, in order to make the code faster.

The second figure plots the actual real-time predictions for every second in a greyscale image. Since the neural network outputs the probabilities for every of the five sleep stages in every second, this plot shows all the predictions in one image. This plot enables the researcher to analyze, whether the AI was biased towards one or two sleep stages, e. g. predicting N3 sleep most of the night, as can be seen in figure 3.18.

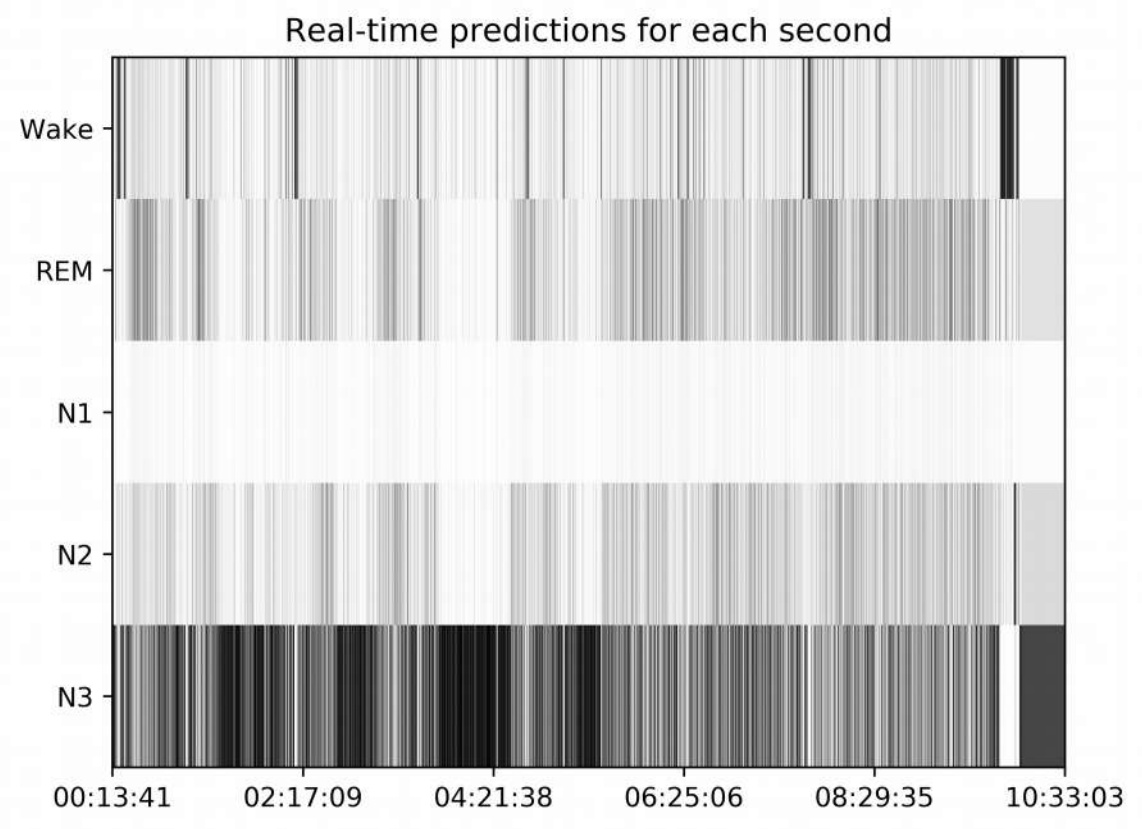


Figure 3.18: This plot shows the real-time sleep stage predictions of the Traumschreiber system. During each second of the recording, the neural network outputs were logged. Black means, that the corresponding sleep stage was very likely at this second, white means the opposite (gray in between). One can see, that the neural network was biased towards predicting N3 sleep in this recording.

The last figure consists of six subplots. In the first column, first plot, the predictions for each sleep stage averaged over 20 minutes (sliding window) are displayed as they were calculated in real-time. Below one can see the resulting hypnogram, depicting which sleep stage was most likely during each second. In the second column, the same is shown, but this time including the additive bias correction term, which takes the subject-individual sleep stage distributions of previous nights into consideration and tries to adjust the real-time predicted sleep stages accordingly. For example, if the neural network is prone to overestimate N3 sleep for an individual subject, in the following nights the bias term tries to correct this. The last column also shows the neural network predictions of the same night, but is post-optimized. This means, that, like in the second column, the original predictions are corrected using an additive bias term. The difference is, that in the third column knowledge is used, which is not available during the time of recording, namely the overall sleep stage distribution of the recorded night (thus post-optimized). An additive bias correction term is estimated based on this distribution, which shifts the sleep stage probabilities up or down. See figure 3.19 for an example.

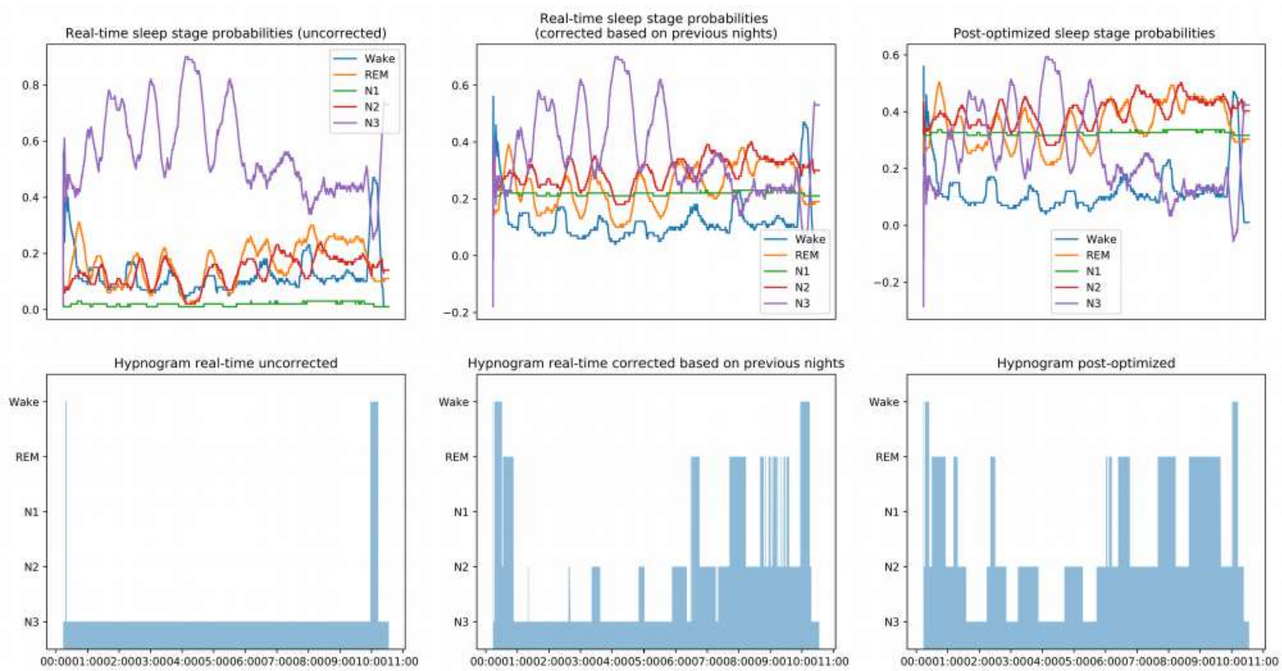


Figure 3.19: The predictions of the same night as depicted in figure 3.18, but averaged using a 20 minute sliding window. The left two plots show the averaged 20 minute probabilities (top: for each sleep stage individually, bottom: displaying only the sleep stage with the highest probability). The two plots in the center show the same predictions, but with an additive bias, which was calculated based on previous recordings of this subject. In the right plots, the nightly predictions were post-optimized, by shifting them vertically such that the distribution of sleep stages comes close to the theoretical average of healthy human sleep, removing most of the real-time prediction bias towards N3 sleep. Typical sleep cycles can be seen in the post-optimized calculation.

3.2 Evaluation

After developing the system, it was evaluated in several studies.

3.2.1 First study: Sleep laboratory validation

This study compared the simultaneous recordings of six subjects with both the Traumschreiber system and the ALICE system, a commercial polysomnographic system, which is used in clinical and research sleep laboratories worldwide. In total, 189 30-second epochs of sleep data were compared to each other, i. e. on average 31.5±0.84 epochs of each subject. The epochs were selected based on their time of recording, with all epochs sampled at 15 minute intervals. Screenshots of all epochs as well as the complete evaluation table can be found in the online appendix. See figures 3.20 and 3.21 for one exemplary comparison.

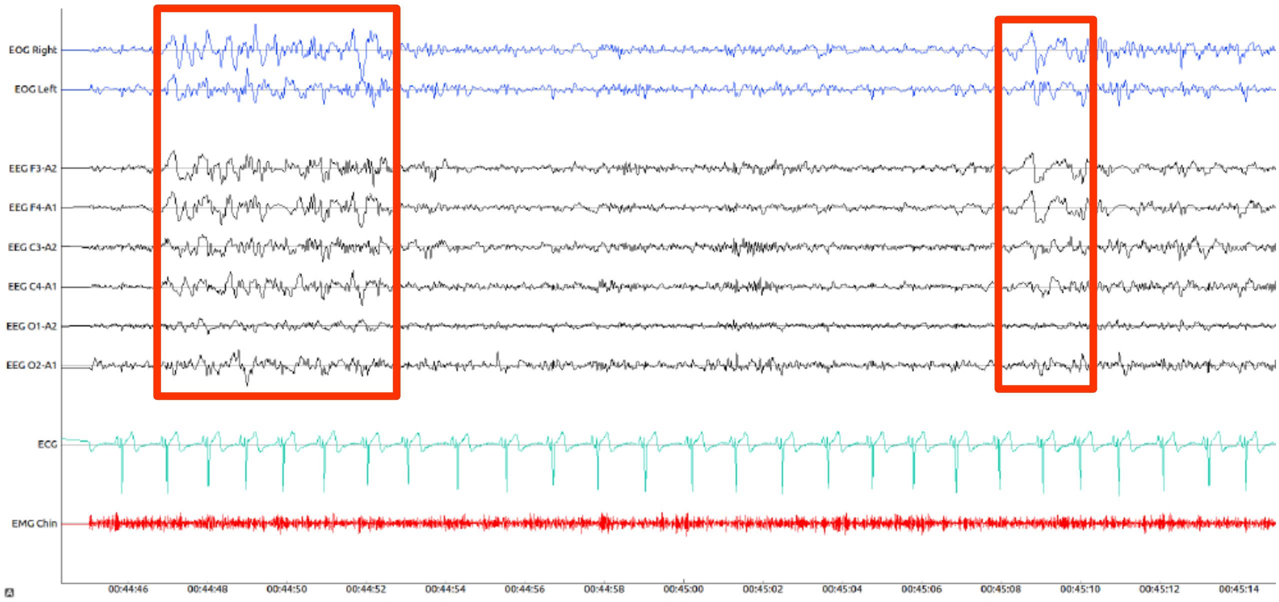


Figure 3.20: This plot shows 30 seconds of polysomnographic data recorded with the commercially available ALICE system, which is used in clinical and research sleep laboratories worldwide.

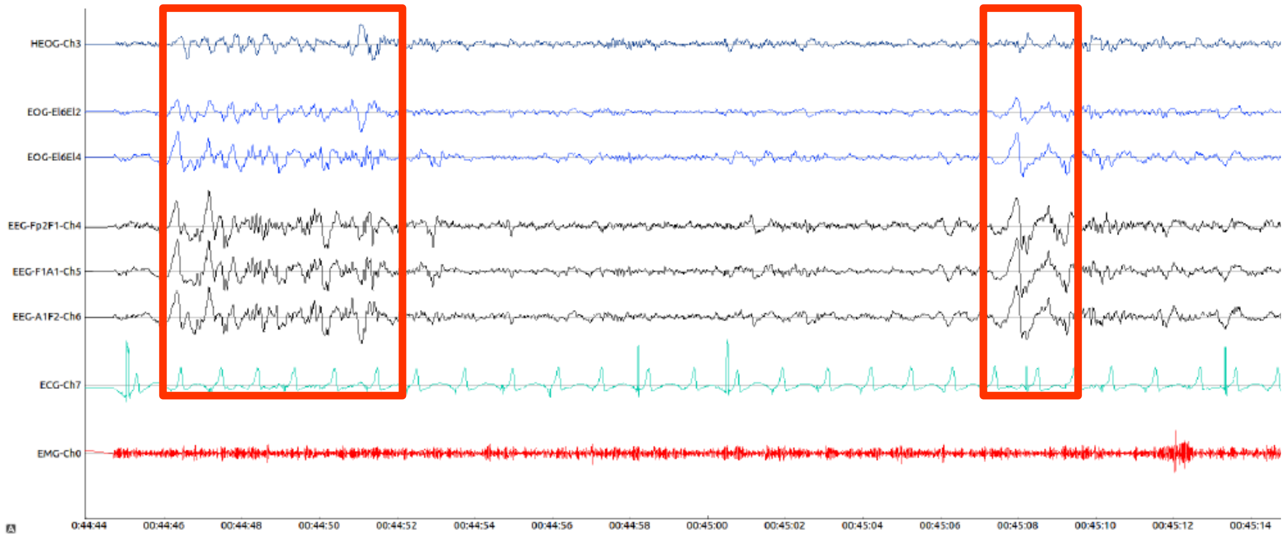


Figure 3.21: This plot shows (approximately) the same 30 seconds of data, recorded with the Traumschreiber system. The red boxes are inserted in order to help the reader to compare both figures.

3.2.1.1 Direct comparison of both systems

In general, both systems recorded very similar data. All sleep graphoelements, except central-occipital sleep spindles and alpha EEG activity, could be detected in both systems in nearly all corresponding epochs, i. e. if one system showed slow waves, the other systems showed slow waves, too.

	Slow waves	Sleep spindles	K-complexes	Low amplitude, mixed frequency EEG	REMs	Heart beats	Arousals	EEG alpha band activity	SEMs	Eye blinks	Vertex waves
	… are visible in __ of the 189 analyzed epochs …										
only in the ALICE system	1	26	3	2	0	0	1	3	0	0	0
in the same epoch in both systems	36	58	53	95	13	189	19	3	1	2	2
only in the Traumschreiber system	0	1	0	0	3	0	0	0	0	0	0

Table 3.3: This comparison shows, which polysomnographic graphoelements could be found in how many corresponding epochs of both systems.

Slow waves could be detected in 36 of the epochs in both systems simultaneously and in one epoch only in the ALICE system. Sleep spindles were clearly visible in 58 corresponding epochs of both systems, 26 times only in the ALICE data, and once only in the Traumschreiber data. In nearly all of the 26 cases, in which spindles were only detectable in the ALICE data, these spindles were measured in the central-occipital region, and were not visible in the ALICE frontal EEG channels, too. K-complexes could be detected in 56 epochs, out of which three were only ALICE recordings. A low amplitude and mixed frequency EEG signal was apparent 95 times in both systems simultaneously, and two times only in the ALICE system. Rapid eye movements could be found in 13 of the epochs in both ALICE and Traumschreiber, and three times only in the Traumschreiber system; slow eye movements only once (both systems). The heart beat was clearly visible in all epochs of both systems. Arousals were found 19 times in both and once only in the ALICE system. Alpha wave activity was detected in three corresponding epochs of both systems, and three times only in the ALICE recordings. Vertex waves were found two times in both systems. The EMG often looked completely different, however, it is unclear whether this was due to a bad signal of the ALICE or the Traumschreiber system, since arousals were detectable in the Traumschreiber system’s EMG signal much better than in the ALICE system.

3.2.1.2 Sleep staging comparison

In 175 of the 189 epochs (92.6%), the same sleep stage was scored for both systems based on inspecting the corresponding 30 second windows: 14x wake, 13x REM, 75x N2, 28x N3 and 45x ambivalent, i. e. it remained unclear, whether wake, (tonic) REM, N1 or N2 should be scored, because previous epochs would have been necessary for an exact scoring.

The remaining 14 epochs with different scoring contained 13x ambivalent and 1x REM scoring for the Traumschreiber data, and 11x N2 and 3x wake classifications for the ALICE epochs, i. e. the ALICE system provided some extra clarity regarding the current sleep stage. In most cases, this was due to central-occipital sleep spindle activity.

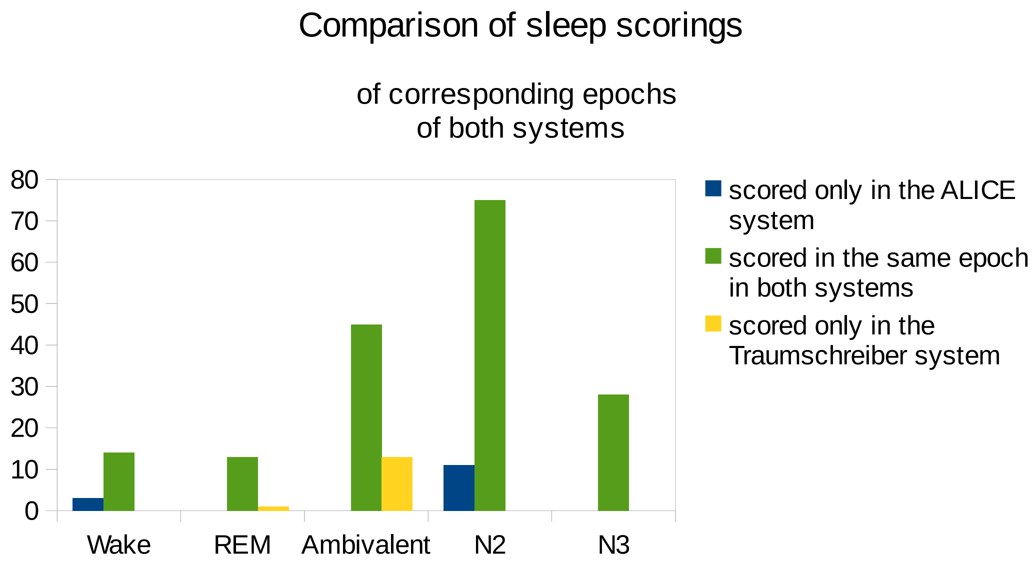


Figure 3.22: The plot shows, how many times the same sleep stage was scored for the 189 corresponding epochs in both systems (green), and in how many cases an epoch was assigned a specific sleep stage in only one of the systems (yellow and blue).

3.2.2 Second study: Home-based prototype study

In this study, a prototype (version 2.2) of the Traumschreiber system was tested at the subjects' homes for the first time, i. e. without a human experimenter at the location of recording. Viewed holistically, the approach to record polysomnographic data during a field study sleep experiment at the subject's home was successful: The Traumschreiber system enabled the naive subjects to carry out the interactive polysomnographic sleep experiment at home, and conducted parts of the experiment on its own. This included playback of acoustic cues at different sound intensities, when the subject was asleep.

3.2.2.1 Data quality

In 14 of the 23 recordings with this prototype version 2.2, data were transmitted from the sleep mask and saved by the minicomputer during the whole night (8 recordings with no signal loss at all, 6 recordings with only a few minutes of signal loss). During the other nights, the data were either not saved or not transmitted throughout the whole night (note that in this (not final) prototype, a lost bluetooth connection was not automatically re-established).

The data quality of the fully recorded nights was analyzed next. The EMG channel showed a good signal in 13 of the 14 nights (93%), with a good signal being defined as 90% or more of the night showing clear data without artifacts. At least one of the two EOG channels (VEOG and HEOG) showed a good signal in 100% of the nights, with both of them having a good signal in 71% of the recordings. In 86% of the nights, at least one of the two EEG channels yielded good signal quality. ECG could be recorded in 86% of the nights in at least one of the two channels. Note that only two EEG channels were recorded instead of three as suggested in the final systems layout, but one ECG channel more.

3.2.2.2 Effect of the system on sleep and dreaming

The total sleep time was nearly the same in both conditions (mean sleep time: 7:56 hours with wearing the sleep mask at night and 8:06 hours without the sleep mask in the control night). The sleep quality was much better in the control night (without the sleep mask) than during the first night of the study. This was the case for both the acoustic stimulation condition (d=2.9, p<0.001) and the condition without stimulation (d=1.5, p<0.01), as can also be seen in figure 3.23.

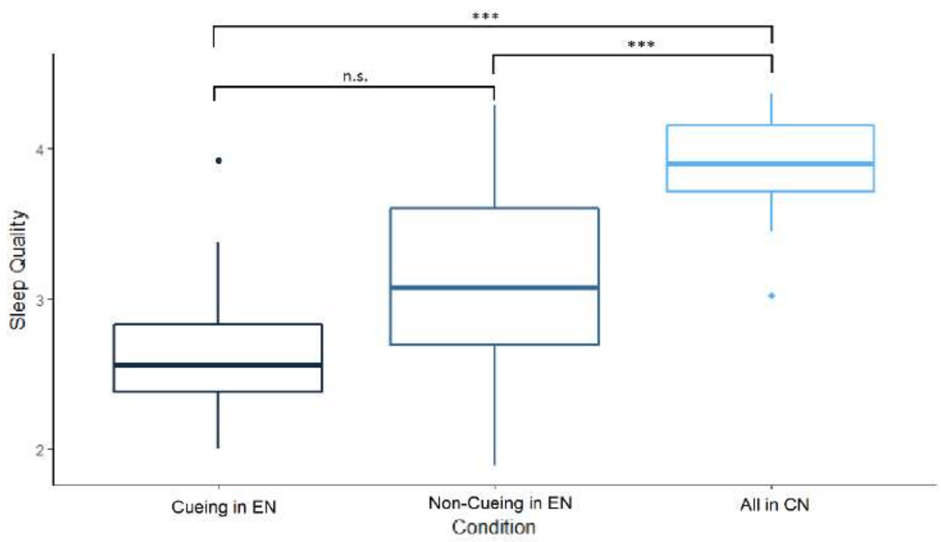


Figure 3.23: This plot depicts the sleep quality in the first night (experimental night, EN), both for the condition with auditory cues throughout the night (left) and without (center), as well as for the control night (CN) without wearing the sleep mask one week later (right). With permission from (Mandt, 2017)

The feeling of recovery was significantly better after the control night, than in the night with the sleep mask (taken both stimulation conditions together; d=0.82, p<0.01). It was, however, not significantly better, if compared to only the no-stimulation night (p=0.08). Please see figure 3.24 for details. The time to fall asleep was significantly shorter during the control night (d=1.1, p<0.01), which was due to the sleep mask, the electrodes, the unusual circumstances or the acoustic stimuli, as the subjects reported (see table 3.4).

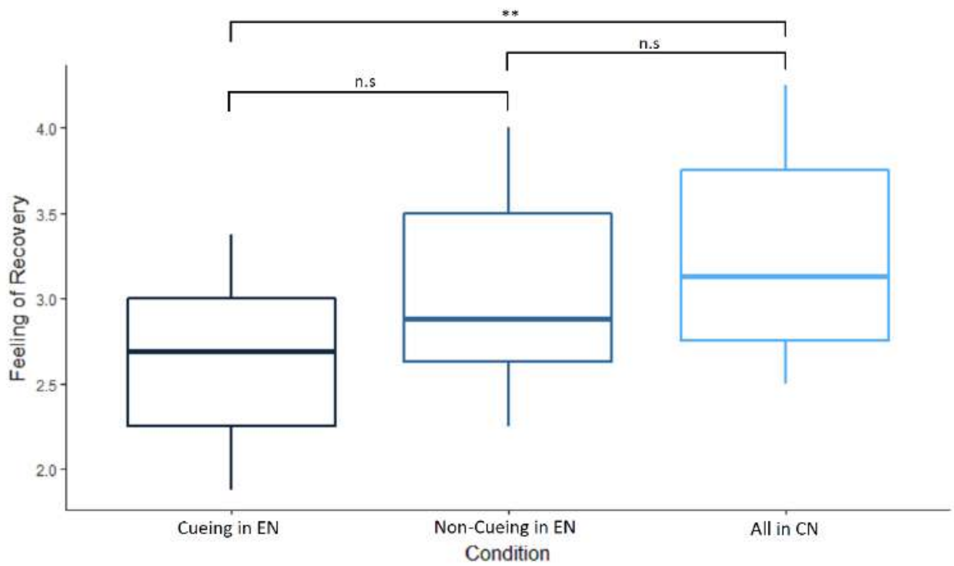


Figure 3.24: This plot shows the feeling of recovery in the morning after the experimental night (EN), for both conditions with cueing (left) and without (center), as well as for the control night (CN) one week later (right). With permission from (Mandt, 2017).

Reasons for not falling asleep	Experimental night		Control night
	Non-Cueing	Cueing	
private reasons/experiences	22.2%	25%	39.1%
sounds	11.1%	33.3%	8.7%
unfamiliar circumstances	22.2%	41.2%	-
sleep mask	44.4%	41.7%	-
exitement	11.1%	8.3%	-
no reason	33.3%	16.7%	60.9%

Table 3.4: The table summarizes the answers of the subjects to the question, what the reasons were for having trouble falling asleep in the beginning of the experimental night (non-cueing and cueing condition) and the control night without wearing the sleep mask. With permission from (Mandt, 2017).

11 of the 14 subjects of the acoustic stimulation condition (78.6%) recalled a dream in the morning, whereas only 4 of the 9 subjects of the condition without stimulation (44.4%) reported a dream. In the control night, 11 of the 23 subjects (47.8%) recalled a dream. The difference between stimulation and no stimulation regarding dream recall was found to be statistically significant with a medium effect size ($d=0.766$, $p=0.034$). The experiment was incorporated into six dreams of the stimulation condition, one dream of the condition without stimulation, and no dream of the control night. The stimuli were not incorporated in any reported dreams.

3.2.2.3 Ease of use and comfort of the sleep mask

All except one subject reported, that they could follow the instructions easily (65.2%) or rather easily (30.4%). Attaching the electrodes to the skin took 3 to 30 minutes (average: 15:02 minutes), and removing them in the morning took 2 to 15 minutes (average: 5:38 minutes). Two subjects had difficulties with attaching the electrodes, and four subjects had problems with removing them in the morning (too sticky electrodes). All other subjects did not report any complications.

The majority of the subjects found the sleep mask not at all disturbing for falling asleep (4.3%) or only slightly disturbing (60.9%), while 34.8% rated it as quite disturbing, with no subject stating it to be extremely disturbing. 52.2% of the subjects were not at all disturbed by the cables of the sleep mask, 26.1% slightly, 17.4% were quite disturbed, and 4.3% felt extremely disturbed. The sleep mask was rated as very comfortable or comfortable by 17.4% of the subjects, as okay by 47.8% of the subjects, as uncomfortable by 30.4% of the subjects, and as very uncomfortable by one subject (4.3%). The material was overall perceived as comfortable. 17.4% of the subjects sweated under the sleep mask.

3.2.2.4 Effect of acoustic stimulation on the occurrence of arousal

On average, each subject of the stimulation group was exposed to 45 acoustic tones of different volume throughout the whole night (one tone every ten minutes). Averaged over all subjects, no arousal was detected after $76.1 \pm 10.4\%$ of the stimulations, and an arousal was found after $23.9 \pm 10.4\%$ of the stimulations. The average stimulus sound volume in case of no arousals was 0.41 ± 0.04, and 0.64 ± 0.08 in case of an arousal. As expected, louder stimuli evoked an arousal significantly more often than quieter stimuli with a large effect size ($d=3.2$, $p<0.001$, see figure 3.25).

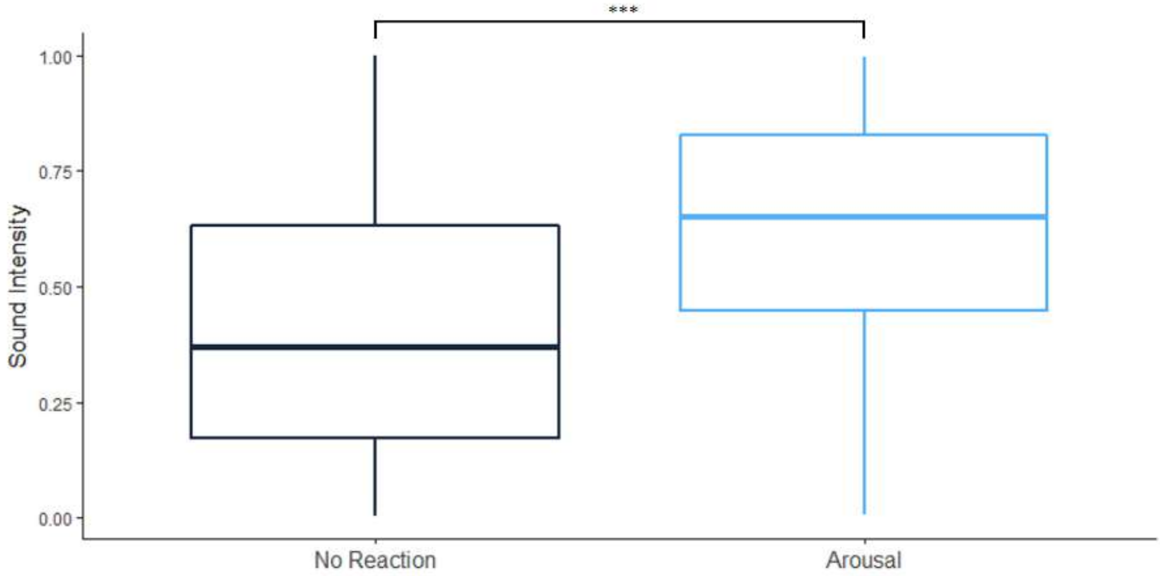


Figure 3.25: The plot shows, that those acoustic stimuli, which lead to arousals, were significantly louder than the cues, which did not provoke an arousal. With permission from (Mandt, 2017).

3.2.2.5 How to further improve the system

The subjects proposed several suggestions, of how the prototype version 2.2 could be improved. Regarding the textile design of the sleep mask, an adjustable elastic head band, a less heavy and smaller battery and a covering cap for the power switch were named.

One of the most important points of how to technically improve the prototype was, that the system should constantly try to reconnect, in case the bluetooth connection gets lost.

Furthermore, the cables should be shortened in order to increase comfort and to minimize the chance of breaking the cable by accident. Several new questions for the self-developed questionnaire about the sleep mask were identified. Moreover, the study showed, that acoustic stimulation should start the earliest 30 minutes after sending the subject to sleep, as the stimuli hindered the subjects to fall asleep. Finally, more practice of the subject in order to ensure good electrode contact was suggested by the experimenter of the study. One way could be to record more nights, and to use the first night as an adaptation and practice night.

All these suggestions were implemented in the final version of the Traumschreiber system.

3.2.3 Third study: Preparing polysomnographic crowd experiments

The goal of this study was to describe all the tasks, which have to be carried out in order to produce one or more entities of the Traumschreiber system for large scale crowd-based experiments. Moreover, a cost analysis was conducted.

3.2.3.1 Description of production work step and cost analysis

Table 3.5 summarizes the work steps based on the experiences from setting up the fourth study of this dissertation. It shows, which tasks have to be conducted and how much time has to be calculated for each work step.

#	**Work step**	**Time**
1	Ordering the components	50 minutes (6 weeks waiting time)
2	Unpacking all the components	4 minutes
3	Cutting the battery holder	3 to 10 minutes
4	Cutting and soldering the electrode cables to the board	14 to 25 minutes
5	Copying the code to the microcontroller and to the BLE chip	2 to 30 minutes
6	Sewing the textile part of the sleep mask and placing the board inside	26 to 90 minutes
7	Copying the Traumschreiber system’s code to the micro-SD card of the experiment station	1 minute (1 hour waiting time)
8	Putting the parts of the experiment station together and into a box	6 minutes
9	Initial charging of the coin cell battery	1 minute (0 to 240 minutes waiting time)
10	Functionality test of the system	3 minutes
	Total production time after the components arrived	*60 to 170 minutes*

Table 3.5: This table gives an overview of the work steps, which are necessary for producing a Traumschreiber system, and the time needed for each step. Please see the following text for further details.

1. **Ordering the components**

   To start producing one or many Traumschreiber systems, the raw materials and components need to be obtained. These include for each Traumschreiber system (costs as ordered in March, 2017):

   <u>For the sleep mask</u>

   ◦ A PCB assembled with all the electric components and a battery holder. This is the most costly and most complicated part to order. It can be produced by several Western and Chinese factories. As became clear during development, the Chinese factories offer the same products at much lower prices than the German factories. However, it is hard to know beforehand, how good the quality of the produced goods is. Two Chinese factories have been tried out during the development of the Traumschreiber system. While PCBCART produced boards of excellent quality at two separate orders and offered a very good service, ALLPCB produced boards of which only 10% worked properly. This might have been due to bad luck.

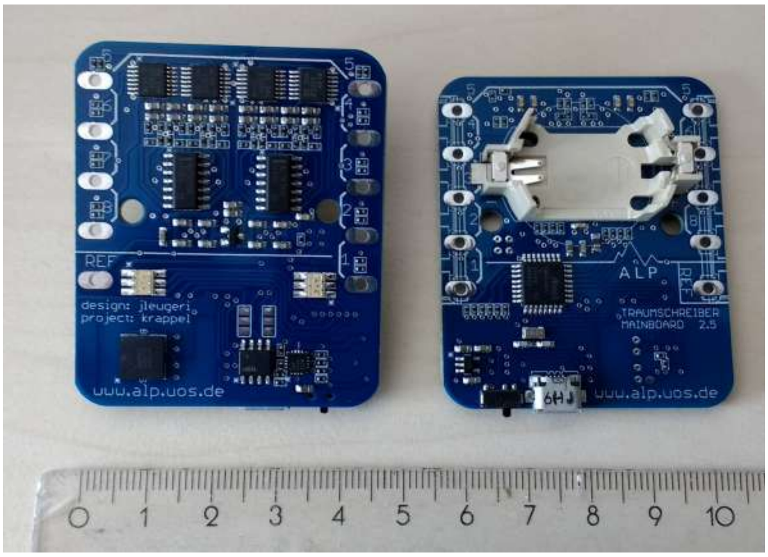


Figure 3.26: Back and front side of the equipped PCB, as it arrives from the factory.

For obtaining a quote and for ordering an assembled PCB, one has to submit a bill of materials (BOM) file and a PCB design file (in optimal case a GERBER RS-274X file) to the factory, e. g. at www.pcbcart.com/assembly, and has to specify several technical details. See the online appendix for the files and details of the last order of the Traumschreiber board.

The costs decrease dramatically the larger the order gets (assembly costs for an order quantity of 10 boards: 56.45 USD per board, for an order quantity of 100 boards: 11.56 USD per board; component costs per board 51.18 USD (order quantity of 10) versus 43.27 USD (order quantity of 100); minimum order quantity costs per board: 14.37 USD (order of 10 boards) v 0.00 USD (order of 100 boards).

The last order during development was placed at PCBCART and contained 20 boards. The total PCB costs were 244 USD, assembly costs (including tooling costs and electric components) 1630 USD, shipping cost from China to Germany 53 USD. Thus, the total costs at this rather small order quantity (20 boards) were 1858 *EUR*, or 93 *EUR* per board.

- 10 electrode cables with 4mm snap-on connectors. These can be ordered in packages of three (red, black, white) at www.olimex.com. This costs 40 EUR for small quantities (12 cables), for large quantities (>49 pcs.) 36 EUR. It is possible to import these cables directly from China at much lower costs, but only in larger quantities. For example, a trader (Lilian Zhang) at www.made-in-china.com offered to produce the same type of cable for 0.92 USD per piece in a quantity of 1000, or 1.2 USD per piece for 500 cables (see invoice in the appendix). A sample order showed, that the Chinese cables were of the same quality as the Bulgarian ones from Olimex.

  Note that in principle any type of EEG/ECG cable can be used – one is not restricted to 4mm snap connectors. Thus, if one would like to record data from occipital regions using reusable cup electrodes, these can be connected to the board as well.

- A 3.6V 180mAh coin cell battery LIR2477, which can be ordered at www.conrad.de for around 5 EUR.

- Molded sleep mask, e. g. model SODIAL, which is imported from China and can be ordered at www.amazon.de for around 2 EUR. As was experienced during development, more expensive sleep masks are not necessarily better, as they are often not molded.
- Piece of cloth (cotton, at least 30x12 cm), hook and loop fastener, and eyelets (all needed for the little pocket containing the electronics). Can be bought at any local textile shop for around 5 EUR.

For the experiment station

- Raspberry Pi 3 Model B including a micro-SD card (class10, with at least 8 GB storage capacity), a case, a power supply (5.1V, 2.5A), and 2 heat sinks. Can be ordered in a bundle, e. g. at www.rasppishop.de for the German market. Costs: around 60 EUR per bundle.
- A pair of USB speakers “Trust Leto 2.0 USB” or similar. Can be ordered at www.amazon.de for around 9 EUR.
- CSL external USB sound card or similar. Can be ordered at www.amazon.de for around 6 EUR.
- Carton box (recommended size appr. 10x17x27 cm) for packing everything. Can be ordered at www.memolife.de for around 2 EUR.

For the recordings

- Covidien Kendall H135SG single use sticky electrodes. 10 electrodes are needed for each recording. Can be ordered at www.ternimed.de for around 7 EUR per package of 50 electrodes, costs reduce slightly for larger quantities.

For the 20 complete Traumschreiber systems, which were produced during the last development iteration in March 2017, the total costs per Traumschreiber system were around 138 EUR for the sleep mask components, 77 EUR for the experiment station, and 7 EUR for recording up to 5 nights with each sleep mask – totaling to 222 EUR for the complete system.

Figure 3.27: All parts of 20 complete Traumschreiber systems.

Note, that if an Android app was used instead of the Raspberry Pi, the costs would decrease by 34% to 147 EUR (assuming that the subjects have an Android smartphone or tablet with BLE). Also note, that the electrode cables make up 26% of the costs of the sleep mask, which is ridiculously high, especially since one half of the cable is thrown away during production anyway (see below). Moreover, note that the production costs depicted here are for a quantity of 20 systems; larger production quantities decrease the costs per item substantially.

Furthermore note, that the depicted costs are the initial fixed costs – after production, for every recording only the single use sticky electrodes have to be bought (variable costs), which is around 1.50 EUR per night (assuming that the sleep masks don’t break). Writing of initial fixed costs, the experimenter has to be aware, that some tools are needed for production. These are: eyelet pliers (can be bought for less than 10 EUR), a soldering iron and solder (can be bought for less than 40 EUR), a pair of scissors and a ruler (less than 5 EUR), a scalpel (less than 5 EUR), a programming tool for the microcontroller (e. g. the ATMEL AVR Programmer for less than 40 EUR) and a USB-to-UART adapter for resetting and programming the bluetooth chip (less than 5 EUR). Moreover, for larger sleep mask production quantities, a sewing machine is recommended (around 60 EUR).

The time, which is needed for ordering all the components, is independent of the order quantity. If it takes 5 minutes to order the parts in each of the six online shops plus 20 minutes for ordering the assembled PCB (as this is more complicated), about 50 minutes are needed for ordering.

Based on the experiences of several development iterations, the delivery times can be estimated as follows: The parts of the system, which are bought in European online shops (including Olimex from Bulgaria), usually arrive within a few working days. The textile sleep mask, which is imported from China via Amazon, needs around 3-4 weeks to arrive in Europe. The assembled PCBs are delivered around 6 weeks after the order is placed.

2. **Unpacking the parts**

   Once the orders have arrived from the different online shops, they need to be unpacked. Especially for larger production quantities, the time needed for this step should not be underestimated. Based on the experiences during development, unpacking all the components for one Traumschreiber system takes about four minutes.

3. **Cutting the battery holder**

   Unfortunately, the battery holder, which is assembled to the PCB, does not fit the actually used coin cell battery perfectly. This is due to the unusual height of the coin cell, for which no perfectly matching battery holder could be found so far. This means, that the battery holder has to be cut using a scalpel, so that the coin cell battery fits in it. See figures 3.28 and 3.29 for a comparison of how the battery holder looks like before and after cutting.

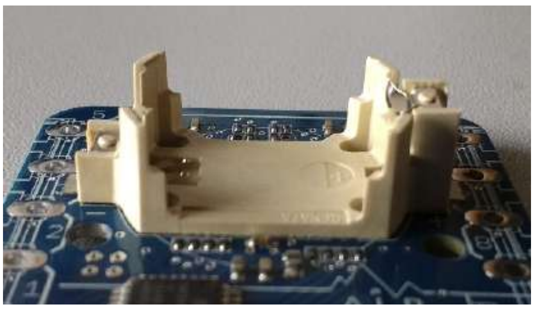


Figure 3.28: The battery holder before cutting.

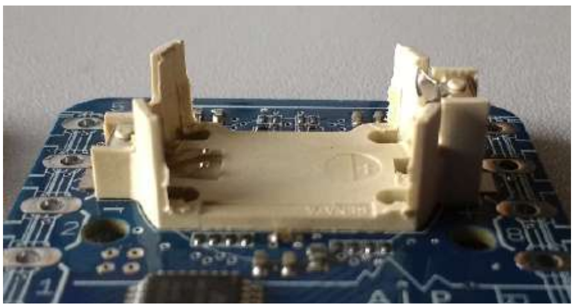


Figure 3.29: The battery holder after cutting.

Cutting the battery holder to a fitting size took on average 5:28 minutes. Noteworthy is that production times speed up due to gaining more practical experience from 9:28 minutes to 2:52, i. e. for larger projects, one can calculate with 3 minutes for this step. The whole step, however, will be unnecessary if a perfectly fitting battery holder is found.

4. **Cutting and soldering the electrode cables to the board**

The ordered cables are 90 cm long and have a 1.5 mm DIN connector at the one end and a 4 mm snap connector at the other end. They need to be shortened, so that no unnecessary long cables hang in the face of the subject. Using a standard pair of scissors, the cables are cut in two pieces of a specified length. The following cable lengths have been found to be useful (measured from the cable end of the 4mm snap connector):

- EMG cables (electrodes 0 and 1): 23 cm and 18.5 cm, preferably use red snap connectors
- EOG cables (electrodes 2, 3 and 4): 10.5 cm, 12.5 cm and 14 cm, preferably use white snap connectors
- EEG cables (electrodes 5, 6 and 7): 8 cm, 19 cm and 20 cm, preferably use black snap connectors
- ECG cable (electrode 8): 62 cm, preferably use red snap connector
- Ground cable (electrode 9): 11.5 cm, preferably use red snap connector.

The second part of the cable (with the 1.5 mm DIN connector) can be thrown away.

Next, the insulation of the cables has to be stripped off at the open end of the cable, and it has to be soldered to the PCB. The cables should be soldered on the side of the battery. With the on-off-switch marking the bottom side of the PCB, the EMG cables are suggested to be soldered clockwise to the bottom left holes, the EOG cables clockwise to the top left, the EEG cables clockwise to the top right, and the other two cables to the bottom right side of the board. See the video in the online appendix (screenshot in figure 3.30) for a detailed illustration of the complete soldering process.

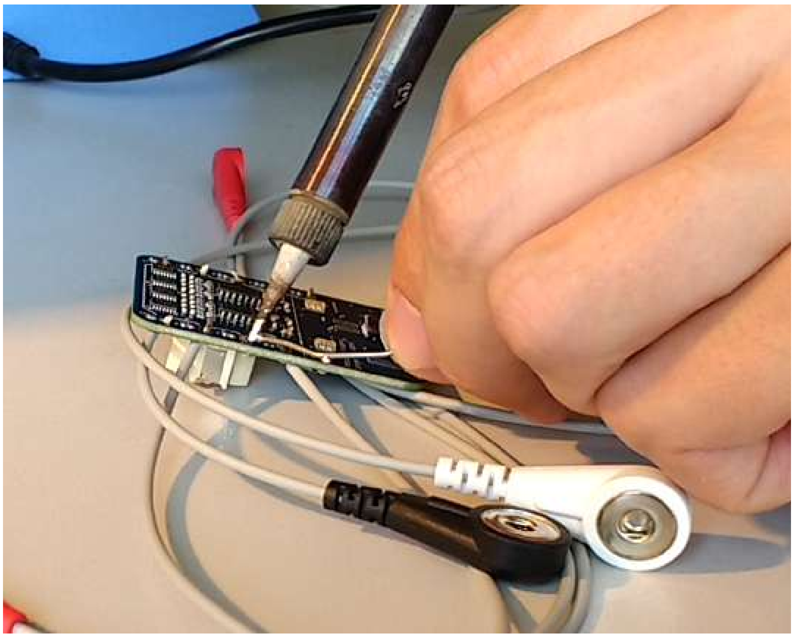


Figure 3.30: Screenshot of the video tutorial, showing how to solder the cables to the Traumschreiber board.

Cutting the cables to the correct sizes takes around 2 minutes per Traumschreiber, removing the isolation 4 minutes (decreasing to 3 minutes after gaining practical experience), and soldering the cables to the board takes a beginner about 19 minutes, decreasing to around 9 minutes after a few soldered boards.

5. **Copying the code to the microcontroller and to the BLE chip**

The microcontroller and the bluetooth chip are delivered without any useful code on them. Thus, they need to be programmed, i. e. the latest development version of the code has to be copied to them.

A programming tool is necessary for this (Atmel AVR programmer for the microcontroller, a USB-to-UART adapter for the bluetooth chip). After connecting the correct pins of the microcontroller or bluetooth chip with the programming tool, the code can be transferred using the software of the programming tool. Further technical details for this step are supplied in the online appendix.

Programming the bluetooth chip took 1:23 minutes on average during the last development iteration. The microcontroller could be programmed in 0:23 minutes on average. Thus, in total about 2 minutes are needed to program both chips. If the procedure is new to the producer, this work step takes considerably longer.

6. Sewing the textile part of the sleep mask

It has been found to be useful to enable the subject to look through the sleep mask, e. g. for putting on the sleep mask using a mirror and for filling out questionnaires during the night. For this, a 4x5 cm rectangular hole needs to be cut into the right eye of the sleep mask and is sewn up.

If the length of the elastic headband of molded sleep mask is not adjustable, it is recommended to cut it and to add a hook and loop fastener to it. The same is the case, if the elastic headband is too short.

Most importantly, the electronic parts need to be integrated into the textile of the sleep mask in order to have a comfortable, safe device. Sewing a tiny bag, in which the electronics are stored and which is then fixated on a molded sleep mask, has been identified as the best solution for this.

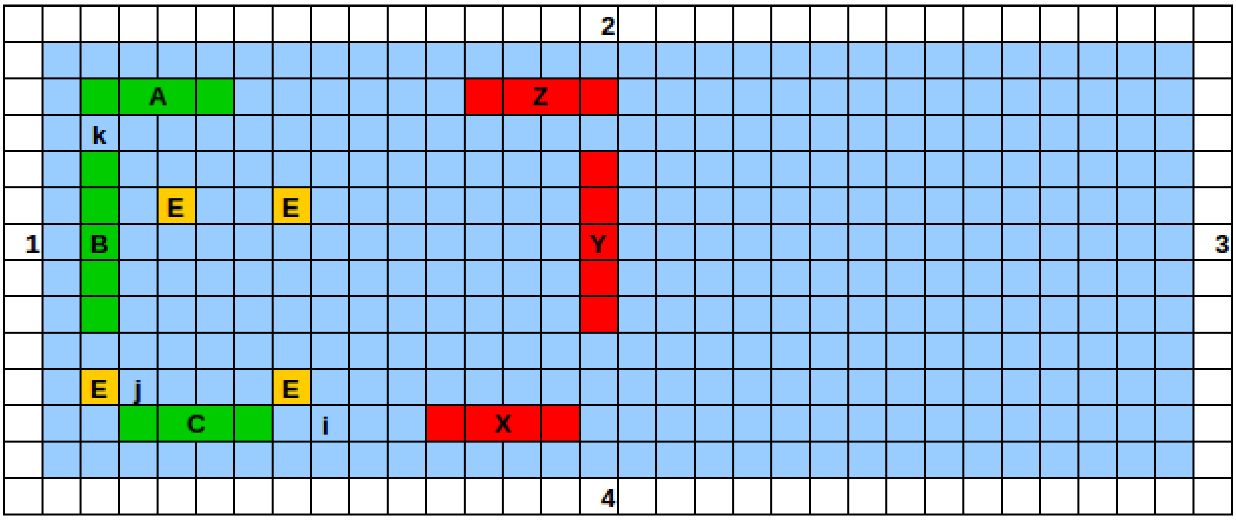


Figure 3.31: Sewing pattern for the sleep mask bag. Each box is 1 cm long. Blue: piece of cotton. Green: hook fastener. Red: loop fastener. Yellow: eyelets.

In order to produce such a tiny bag, a 30x12 cm piece of cloth (cotton, not synthetic, as synthetic material might lead to electrical shorts) has to be equipped with hook and loop fasteners (A, C, X, Y, Z) according to the sewing pattern in figure 3.31. Next, the side 3 is put onto the side 1, with the hook and loop fasteners being inside. The sides 1, 2 and 4 are sewn together, so that only a small (2 cm) hole is left open near the place, where the hook fastener B later is fixated. Next, the bag is turned inside out, and the hook fastener B is sewn onto the cloth, closing the hole. Then the two eyelets near the hook fastener C are used to connect the bag with the molded sleep mask, with both eyelets being fixated in the middle of the top edge of the left eye of the sleep mask. The two other eyelets in the middle are, in a next step, also used connect the bag with the sleep mask. Through these holes, the light of the LEDs will shine into the eyes of the subject. Thus, it is important, that the holes are exactly 3 cm apart from each other. Next, the PCB is placed inside the (open) bag and the cables are used to hold it in position: the electrode cables 4, 7 and 9 are fixated with yarn at fixation point i, the electrode cables 3, 5, and 6 at fixation point j, and the electrode cables 0, 1, 2 and 8 at fixation point k. Finally, the bag is closed. See the video in the online appendix (screenshot in figure 3.32) for a detailed illustration of the complete sewing process.

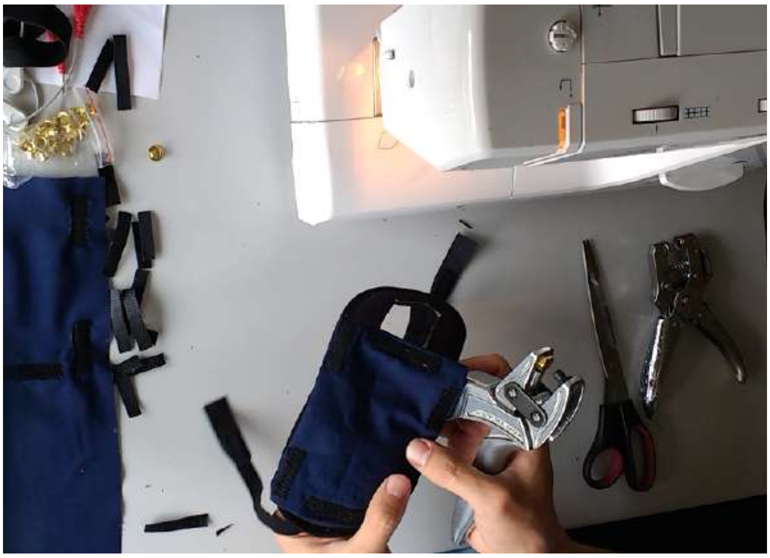


Figure 3.32: Screenshot of the video tutorial, which shows how to create the textile part of the sleep mask.

The whole work step can be conducted in around 26 minutes. However, when inexperienced with the sewing machine and as long as the single sewing steps are unclear, this procedure may take much longer (up to 1.5 hours).

7. **Copying the Traumschreiber software to the micro-SD card of the experiment station**

   For installing the Traumschreiber software, it is recommended to copy a complete image of the latest development version to the micro-SD card (see online appendix). This is probably the most convenient way, ensuring that no python dependencies are forgotten or have to be installed manually over network.

For copying the content to the micro-SD card, the micro-SD card has to be inserted into the card reader of a computer, and the *dd* command has to be used inside a Linux terminal (equivalent tools exist for Windows and Mac as well). The image file is then transferred to the micro-SD card, after which the software is ready to use.

Inserting a micro-SD card into the card reader and executing the dd command takes less than a minute. However, it might take up to one hour until the content has been copied to the micro-SD card.

8. **Putting all the parts of the system together and into a box**

   The parts of the experiment station are delivered separately and maybe even from different online shops. Thus, they need to be put together.

   The Raspberry Pi has to be equipped with the two heat sinks, which just need to be stuck onto the two chips (black "squares") on the top side of the Raspberry Pi. Next, the Raspberry Pi is placed inside the plastic case, and the USB sound card is plugged into any of the four USB ports. The speakers are then connected to both the USB sound card and another USB port, and the power plug is connected to the Raspberry Pi. This ready-to-use experiment station is then placed into the carton box, together with the sleep mask, a package of electrodes, and possibly questionnaires and other paperwork.

   Putting the pieces of the Raspberry Pi together takes around 2 minutes. Folding the box and putting everything inside takes another 4 minutes.

9. **Initial charging of the coin cell battery**

   Even though the coin cell batteries for the sleep mask should arrive fully charged, this is not always the case. Of the 25 batteries tested, four were completely discharged (0.0 V), one was nearly empty (1.11 V), and the other 20 batteries were completely charged (around 4 V). Thus, time should be calculated for an initial charging of the batteries, which can last up to four hours, if the battery is empty. For charging, the micro-USB power adapter of the Raspberry Pi has to be connected to the sleep mask and a power plug, which takes less than a minute.

10. **Functionality test of the system**

    Once everything is set up, the system can be tested by switching on the Raspberry

Pi minicomputer and waiting for the instruction from the USB speakers to switch on the sleep mask. If everything is in order, the Traumschreiber system will detect the sleep mask after a few seconds and it will start saving the data to the micro-SD card every 30 seconds (folder: data/databases). This step takes about 3 minutes in total.

Summed up, any interested scientist should be able to produce a complete Traumschreiber system following these 10 steps. In the beginning it will take longer, but once the procedure is clear, and especially if several systems are produced in bulk, a single person should be able to produce one system in an hour of time – as was the case for the author of this dissertation.

3.2.3.2 Programming the experiment

Once sufficiently many experiment boxes have been produced, the next step is to program the experiment. In order to adjust the experimental protocol to the exact research question in mind, in simple cases, only some parameters have to be changed. In more complex experimental paradigms, the actual Python scripts have to be adapted.

As was described in chapter *3.1.2.2.1 Autonomous, easy to program experiments*, an xml template exists, which can be modified to reflect the experiment design. It is possible to create several kinds of experiments using this standard procedure. Most likely, the beginning of each experiment will consist of welcoming the subjects, asking them to fill out some questionnaires, to prepare for going to bed, to put on the sleep mask according to the instructions in the tutorial video, and to finally go to bed and switch on the sleep mask. Afterwards, a bio signal calibration might be conducted. What happens afterwards, highly depends on the experimental question being investigated. One could just record a night of undisturbed sleep, present stimuli to the subject in regular time intervals or during specific sleep stages, or ask the subject to fill out further questionnaires during the night or in the morning. Moreover, if the study's design is more complicated, one can modify the python code by programming new xml elements or by training and applying advanced EEG pattern detectors, e. g. for closed-loop slow wave sleep stimulation.

Afterwards, the updated experiment xml files (and eventually any other updated files of the Traumschreiber system) need to be copied to the experiment stations.

Finally, the experiment boxes can be handed out to the subjects or can be sent by post.

3.2.4 Fourth study: A crowd-based polysomnographic experiment

This study was carried out in order to investigate, whether the Traumschreiber system enables inexperienced crowd subjects to perform a polysomnographic sleep and dream experiment at home. Moreover, a parallel study design was tested, i. e. more than 10 subjects recorded data in parallel, decreasing the duration of the data acquisition substantially.

3.2.4.1 General results

13 of the 14 subject, who each collected a Traumschreiber system, recorded three full nights. The 14th subject dropped out due to a technical defect (on-off switch broke). The further analysis will be based on the 13 subjects, who participated in the whole experiment.

In all 39 nights, the Traumschreiber system enabled the naive crowd subjects to conduct the sleep and dream experiment at home. More specifically, it carried out the following tasks:

- It welcomed and instructed the subject in the evening about what to do when.
- It conducted the bio signal calibration procedure.
- It wished good night.
- It woke up the subject multiple times during the night according to the procedure specified in the experiment xml file.
- It carried out real-time sleep staging in every second of each night using a neural network based machine learning approach.
- It requested the subject during the night to fill out the nightly questionnaires.
- It told the subject to continue sleeping.
- It protocoled everything (times, actions, stimuli).

During the short time frame of the experiment – exactly one week –, 39 polysomnographic measurements were conducted, and 207 questionnaires were filled out by the subjects during the night, after they were acoustically stimulated by the Traumschreiber system.

Figure 3.33: Example of the bio signal calibration procedure, which the Traumschreiber system carried out at the beginning of each night. One can clearly see the eye movements, which the Traumschreiber system asked the subject to perform, in the VEOG and HEOG channels.

The subjects were asleep prior to the nightly stimulation in 85.2% of the cases, according to their own self-rating, and overslept 15.6% of the acoustic stimulations. Even though no human experimenter was at the place of recording, the subjects' reliability concerning the nightly questionnaires was high: 98.2% of the nightly questionnaires were filled out at the time of nightly stimulation, as was revealed by the coding number, which the subjects could obtain only during the night.

3.2.4.2 Data quality

While the transmission quality was good in only two of the 13 first nights, this changed to 10 out of 13 nights for the second and third recording of each subject (statistically significant, d=1.569, p=0.001, comparing first and second night).

What were the reasons for bad transmission quality? Two subjects reported in the first night, that the on-off switch of their sleep mask was loose. This fits the transmission plots, which indicate fragmentary data transmission for all three nights of these two subjects (see online appendix). Six of the remaining first night recordings with bad transmission quality show an abrupt loose of signal after a few hours, three times directly following the acoustic stimulation. Two reasons might explain this: either the subject (intentionally or

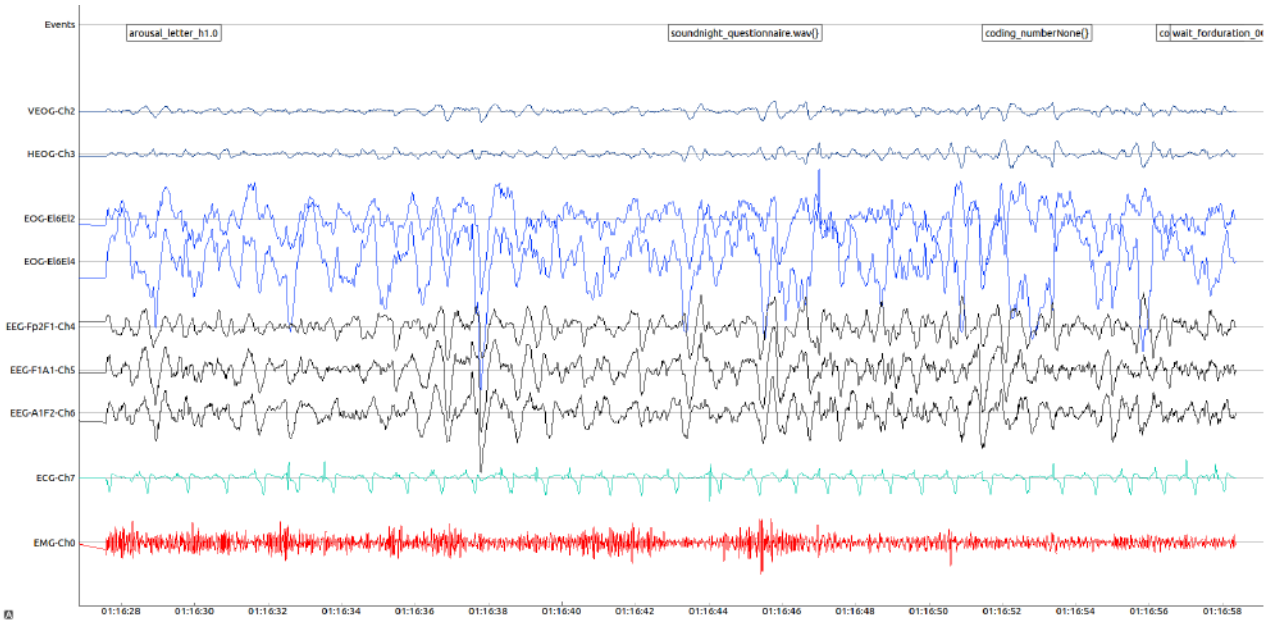


Figure 3.34: Example of a nightly arousal procedure, during which the subject overslept (N3 sleep continues). This matches the nightly questionnaire, which was not filled out by the subject.

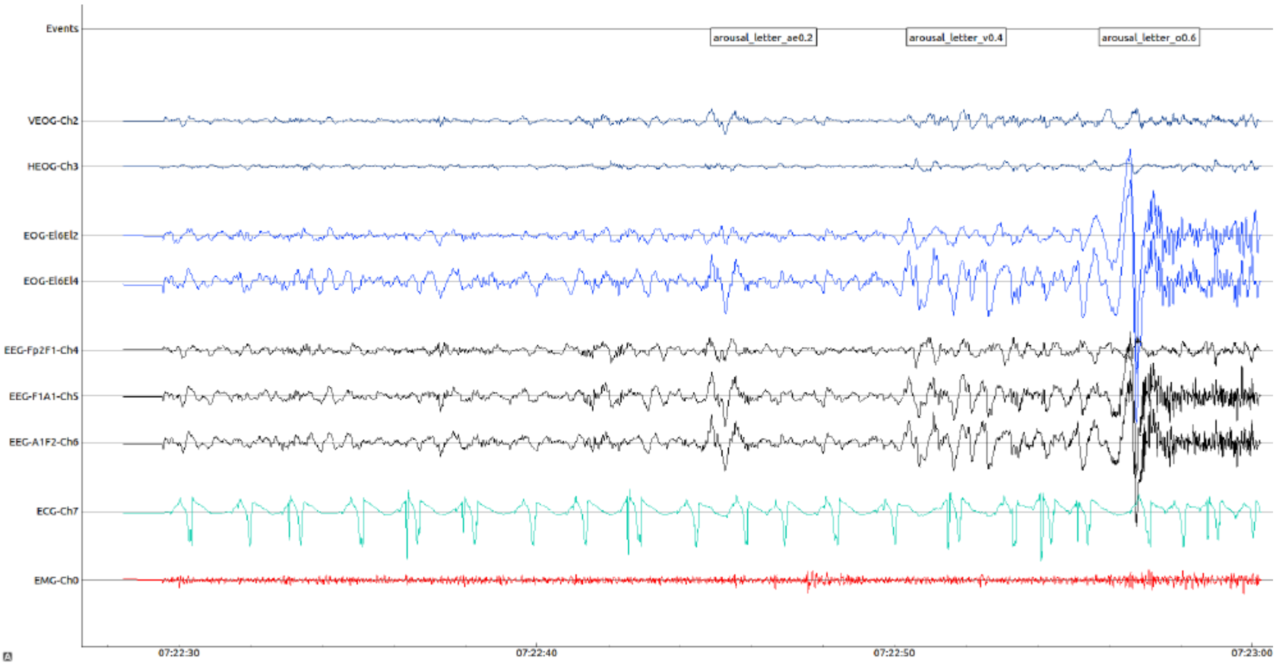


Figure 3.35: Example of a nightly arousal procedure, during which the subject woke up. In the nightly questionnaire, the subject reported to have heard the letter “o” first after waking up, i. e. the subject missed the first two stimuli “ä” and “v”.

unintentionally, consciously or subconsciously) switched off the sleep mask, or the batteries were not fully charged before conducting the experiment the first time (even though all the batteries were checked before handing them out to the subjects – they might have discharged before the first experimental night). In one of the remaining three recordings with bad transmission quality in the first night, no data was received at all, which might be due to the subject having forgotten to switch the sleep mask on, or a completely empty battery. For the two other cases of the first night, it remains unclear, why the data transmission was bad. In the third recording during the second experimental night, which showed poor transmission quality, the data transmission was stopped after about half an hour, suggesting that the battery was not charged. The data of the last recording with bad transmission quality (third night) indicate, that the sleep mask was switched on briefly in the evening for a few seconds, but switched off again. In the morning, shortly before the Raspberry Pi was switched off, again some seconds of data were transmitted, suggesting that the subject accidentally moved the on-off switch in the wrong direction.

Looking at the data quality during the 22 nights with good signal transmission, overall, the quality of the transmitted data was very good. If the first night is considered a training and practice night, and also the two successfully recorded datasets of the first night are excluded, the second and third experimental night (N=20) show a good EEG signal in 19 of the 20 recordings in at least one channel for at least 90% of the night, and in many cases also the other channels showed a good data quality (at least two good EEG channels: 16 recordings, all three EEG channels: 9 recordings). The same is the case for the EMG channel, which had a good data quality in 19 of the 20 recordings. The EOG yielded good data in 18 nights in at least one channel (two channels: 11 recordings), whereas the ECG channel only recorded data of good quality during 15 of the 20 nights. In the two first night recordings with good signal transmission, both the ECG and EMG channel recorded data of good quality in both subjects, whereas only in one of the two recordings a good EEG and EOG signal could be obtained.

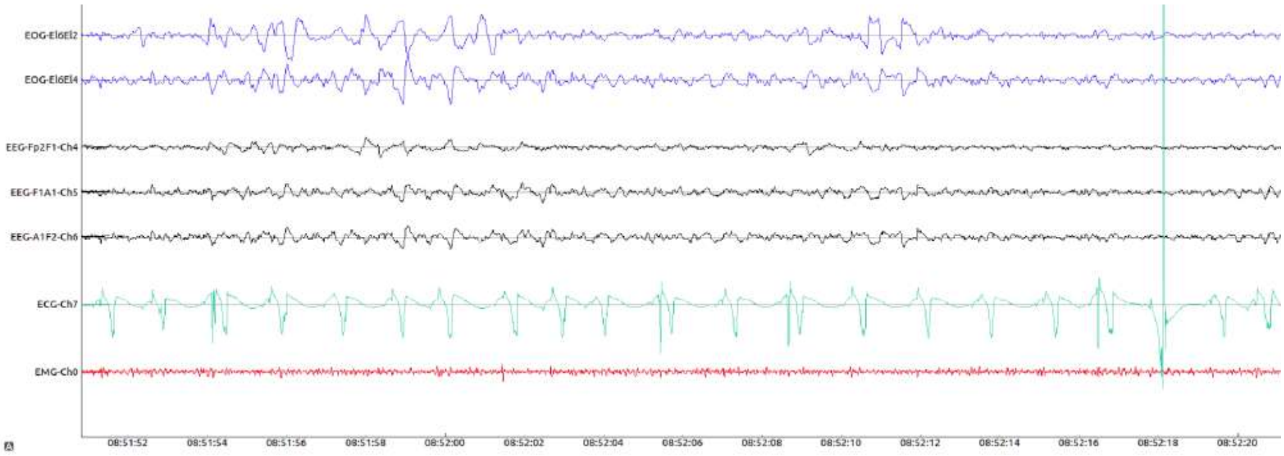


Figure 3.36: Example of REM sleep, recorded autonomously by the Traumschreiber system at the subject's home.

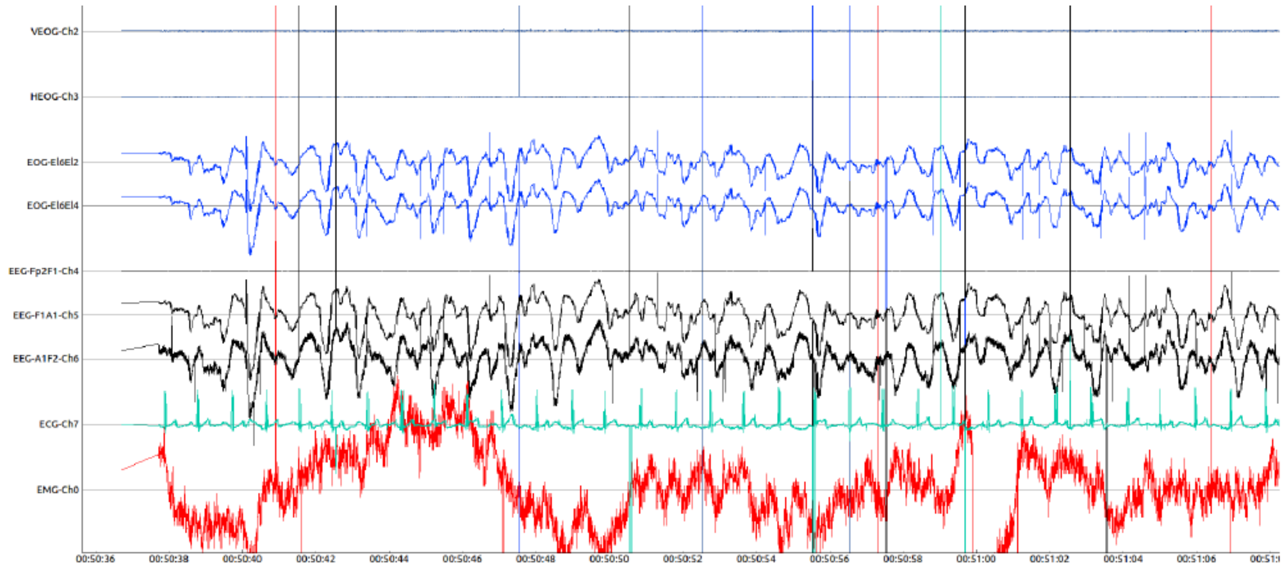


Figure 3.37: Unfiltered, raw data example of a recording with bad quality. The channels 2, 3 and 4 are flat (most likely due to static charges on the electrodes). The EMG (channel 0) shows large drifts, however, these can be filtered out in software using a bandpass filter. Channel 5 and 6 show clearly visible slow waves, with 50 Hz power line noise on channel 6 (can be filtered out in software, too). The ECG shows a clear heart beat. On all channels, spikes are visible, which is unfortunately the case for all recordings of the here used version of the Traumschreiber system, and most likely due to a programming error of the microcontroller and its internal ADC. However, these spikes can be removed using a median filter or a threshold. This bug is removed in the in the latest firmware version, which is attached in the online appendix.

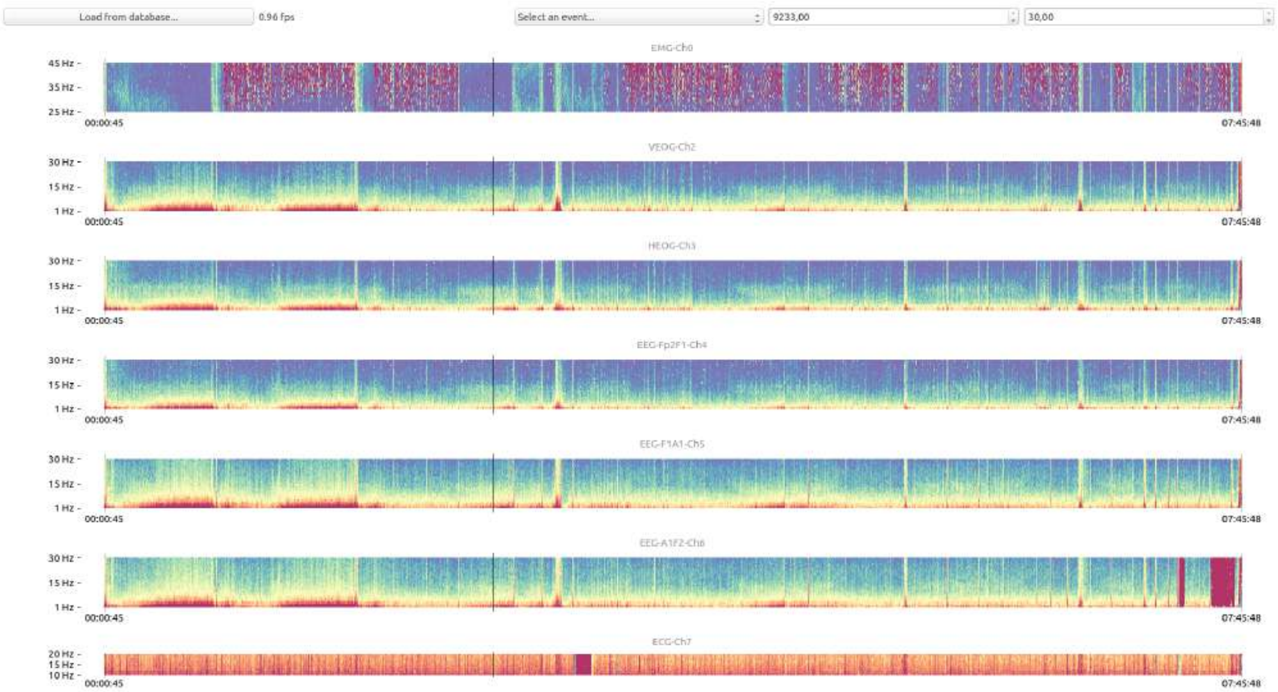


Figure 3.38: Example of a Traumschreiber system recording (time frequency plots of seven channels) with good data quality on all EMG, EEG, EOG and ECG channels in at least 90% of the night. Only one EEG channel (6th row) has a bad data quality at the end of the night (dark red color), as well as the ECG for a few minutes during the night.

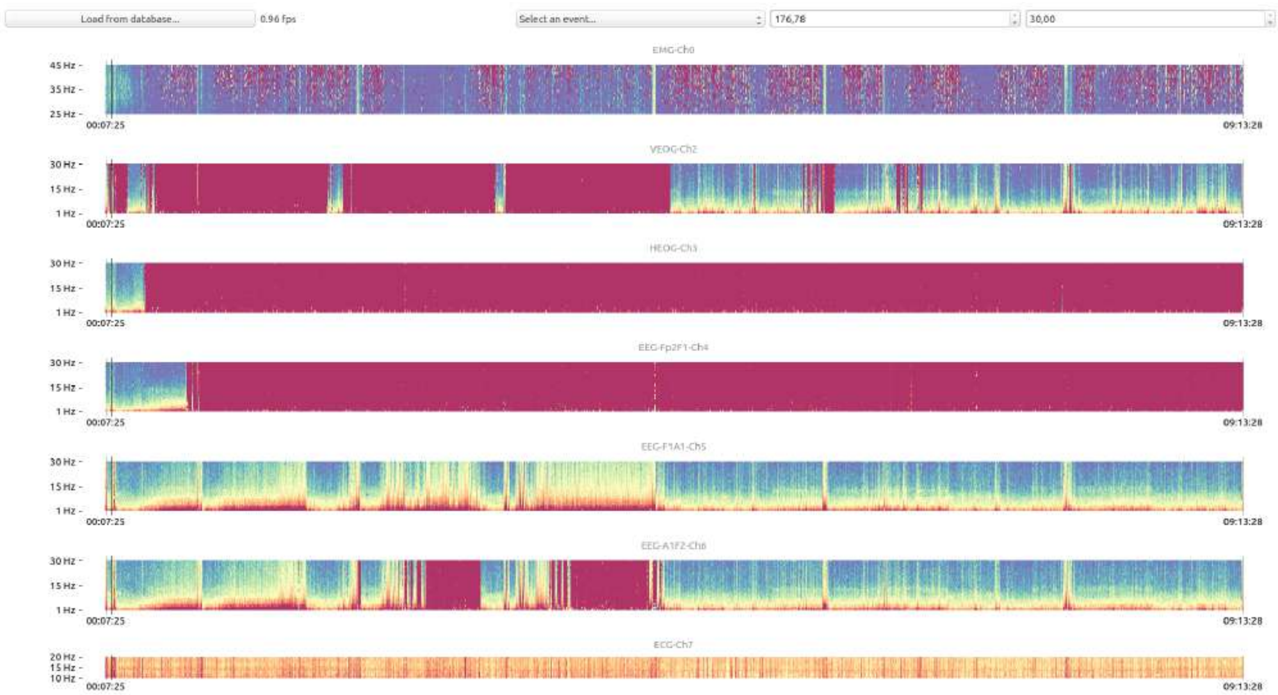


Figure 3.39: Example of a Traumschreiber system recording (time frequency plots of seven channels) with bad data quality in several channels (both EOG channels, two EEG channels).

Taking the transmission success and the data quality together, and excluding the two subjects with loose on-off switches in their sleep masks, 86% of the recordings of the second and third night showed a good EEG signal, 82% a good EOG signal, 86% a good EMG signal and 68% a good ECG signal. Ten of the eleven subjects with intact sleep masks delivered at least one full night of polysomnographic recording during night two or three (EEG, EOG and EMG), six of which recorded both nights two and three successfully. Eight of the eleven subjects recorded whole nights with EEG, EOG, EMG and ECG.

11 subjects, second and third night (N=22)			
Recordings with good EEG signal	Recordings with good EMG signal	Recordings with good EOG signal	Recordings with good ECG signal
86%	**86%**	**82%**	**68%**

Table 3.6: This table shows, how many of the recordings of the eleven subjects with intact sleep masks yielded a good EEG, EMG, EOG, and ECG signal, respectively.

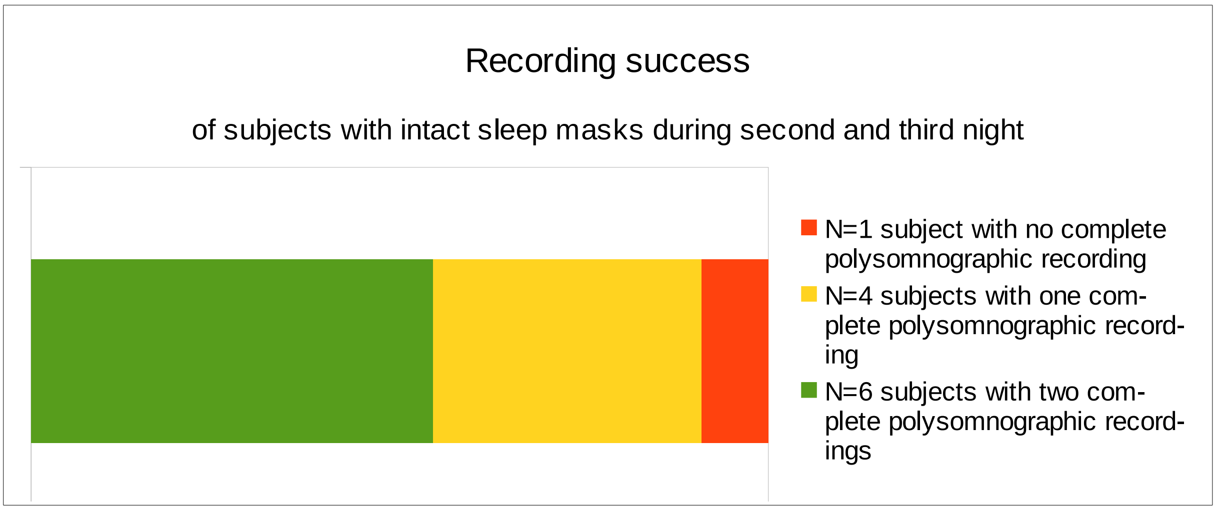


Figure 3.40: This figure shows the recording success during the second and third night of the eleven subjects with intact sleep masks.

There was no significant difference between the second and third night regarding the data quality, except for the ECG channel, which was better during the second night (nine versus six out of ten whole night recordings, d=0.739, p=0.032).

3.2.4.3 *Sleep quality*

The sleep quality (SQ) was not significantly different between the three experimental nights (SQ night 1: 3.03±0.74, night 2: 3.06±0.51, night 3: 3.1±0.7. d(nights 1 and 2)= 0.04, d(nights 1 and 3)= 0.1, d(nights 2 and 3)= 0.08). The feeling of recovery in the morning (FoR) was also not significantly different between the three nights (FoR night 1: 2.89±0.89, night 2: 2.75±0.44, night 3: 2.69±0.84. d(nights 1 and 2)= 0.20, p(nights 1 and 2)=0.661, d(nights 1 and 3)= 0.24, p(nights 1 and 3)=0.618, d(nights 2 and 3)= 0.10).

The sleep quality and the feeling of recovery was worse during this experiment than in the reference groups provided by the questionnaire makers, which is not a surprise as the subjects were woken up on purpose several times during the night (SQ of both reference groups: 3.86±0.81 and 4.05±0.83, FoR: 3.48±0.78, 3.38±0.74).

3.2.4.4 *Ease of use, comfort and durability of the Traumschreiber system*

The subjects found it easy to use the system (putting on electrodes and sleep mask, following the written, auditory and video instructions) – indicated by an average answer to the regarding questions in the sleep mask questionnaire of 1.61±0.60 (scale: 1-easy, 5-difficult). The video instructions were helpful (3.26±0.59, 1-not at all helpful, 4-very helpful). In nearly all nights, the subjects reported that they experienced no technical problems with using the system. In four nights, the subjects reported having switched on the sleep mask too early, so that they had to restart the minicomputer in order to be ready for the bio signal calibration in time. One subject had trouble finding the on-off switch during the first night. Even though the sticky electrodes did not fall off in a single case (as reported by the subjects), it was rather easy to remove them (2.62±1.16, 1-easy, 5-difficult).

The sleep mask and the cables were perceived as slightly disturbing for falling asleep (1.79±0.47, 1-not at all, 4-very disturbing). Overall, the subjects reported, that sleeping with the sleep mask and the sleep mask material was neither comfortable nor uncomfortable (2.92±0.45, 1-very comfortable, 5-very uncomfortable). When asked after the second and third night, whether they slept better or worse than in the previous experiment night, the subjects said that they slept slightly better in the later nights (2.46±0.90, 1-much better, 5-much worse).

No signs of wear were visible after using the sleep mask three nights for each subject, except one ECG cable, which broke during the third night of one subject.

3.2.4.5 Feedback of the subjects to the sleep mask and the experiment

Overall, the feedback was mixed. Some subjects praised the design or reported the sleep mask to be comfortable or at least more comfortable than expected. Other subjects complained about parts of the experiment or made suggestions, how to improve it: Eight of the 13 subjects found the volume of the acoustic stimuli inappropriate (six times too loud, two times too quiet). Four subjects wanted more padding at the ear or the nose part of the sleep mask. Three subjects found it difficult to wash the glue residues off the skin after the sticky electrodes were removed. Two subjects suggested to label the cables better.

3.2.4.6 Automatic sleep scoring algorithm and subject-specific adjustment

All Traumschreiber systems predicted the current sleep stage using the Keras/tensorflow neural network approach during every second of the experiments, in which data was transferred from the sleep mask. Furthermore, the predicted sleep stages were used for steering the experiment and determining the stimulation times. This happened in real-time autonomously, and followed the exact rules given in the experiment xml file.

The quality of the real-time sleep scoring algorithm, i. e. the sleep stages predicted, however, was not sufficient in most recordings, including the ones with very good polysomnographic data quality. Most automatically scored hypnograms were heavily biased towards the N3 sleep stage. Even though the subject-based real-time adjustment of the predictions changed the actually predicted sleep stages in 18 of the 20 second and third nights, this helped improving the prediction quality only in eight nights, and even with the adjustments, the sleep staging quality remained very poor. Figure 3.41 illustrates this exemplarily (see online appendix for further plots).

In the left two plots, one can see that N3 sleep was predicted for basically the whole night, as its probability among all five sleep stages was highest. However, one can also see, that cyclic patterns exist within the N3 sleep predictions, as well as in the other sleep stage predictions. This suggests, that the neural network detected these overall sleep cycle-like patterns, but had a too high prior for the N3 sleep stage.

The plots in the middle column demonstrate, that the predictions were adjusted, since an additive prior was added to the other sleep stages based on the sleep stage distribution of the previous night, and the N3 predictions were shifted downwards. Even though this improved the prediction results slightly, this was not enough in order to deliver a good real-time sleep stage prediction.

In the plots on the right, the priors were calculated automatically retrospectively, knowing the predictions of the actual night and shifting them up or down. Here, one can see a plot, which looks pretty much like a human-scored hypnogram. Classical sleep macro structure elements are visible, e. g. alternating patterns of REM and non-REM sleep, and more deep sleep in the first half of the night than in the second (even though the first and last REM period should probably rather be classified as awake). The slight deviation from the normal 90 minute sleep cycle length is probably because of the six awakenings of the subject due to the auditory stimulation.

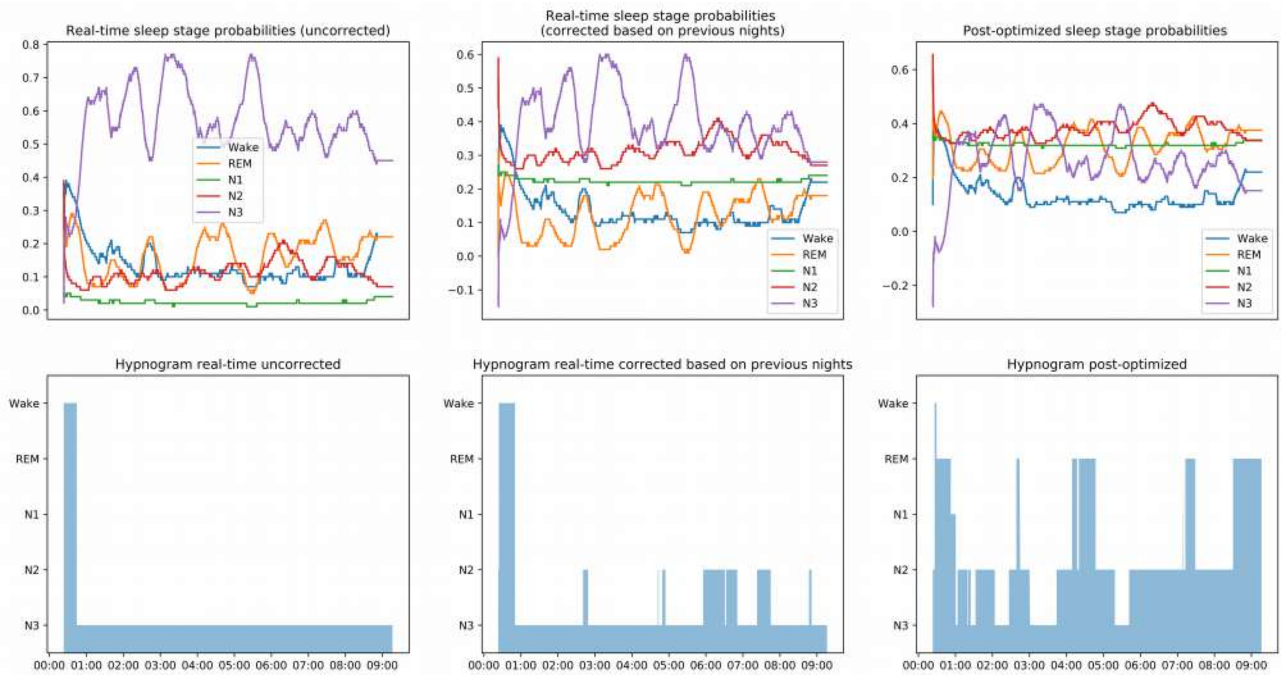


Figure 3.41: Real-time prediction, real-time adjustment, and the automatically post-optimized version.

Summarizing, the Traumschreiber system performed a real-time sleep stage classification using a deep neural network approach. It also adjusted the experiment timings accordingly, and took subject-individual differences into consideration. However, the prediction quality of the simplistic sleep staging algorithm was by far not sufficient and needs further improvement. It has to be kept in mind, though, that only a simplistic sleep staging algorithm was used due to the limited availability of training data, and that, thus, the poor sleep staging performance was to be expected.

3.2.4.7 N3 and REM sleep differences

Ten subjects recalled and rated 18 REM sleep dreams and 14 N3 deep sleep dreams, after they were woken up by the Traumschreiber system during the second or third night (sleep stages scored by a human rater). N3 sleep dreams were attributed significantly stronger by the subjects as being “rather static thoughts, no story” than REM sleep dreams (N3=2.50±1.72, REM=0.67±0.88, d=1.396, p=0.001), and far less “lively or with an action-packed plot” (N3=2.07±2.15, REM=5.78±1.47, d=2.059, p<0.001). REM sleep dreams were described as being much more “entertaining, would like to continue dreaming about it” than N3 sleep dreams (REM=3.89±1.73, N3=1.93±1.94, d=1.074, p=0.006). Even though the REM sleep dreams were rated as being more “bizarre” than the N3 sleep dreams by the subjects, this effect was not statistically significant (REM=2.28±2.08, N3=1.36±1.59, d=0.49, p=0.199). The same is true for the time needed to wake the subject up: There might be a difference between the two sleep stages, but this was not found to be statistically significant (REM=2.74±1.62, N3=3.77±2.07, d=0.558, p=0.064).

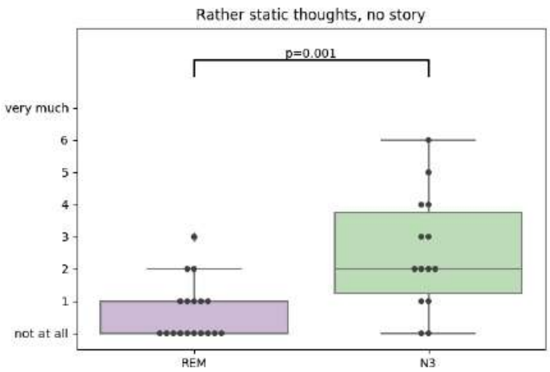


Figure 3.42: Difference between REM and N3 sleep dreams, regarding in how far they were perceived as rather static thoughts with no story.

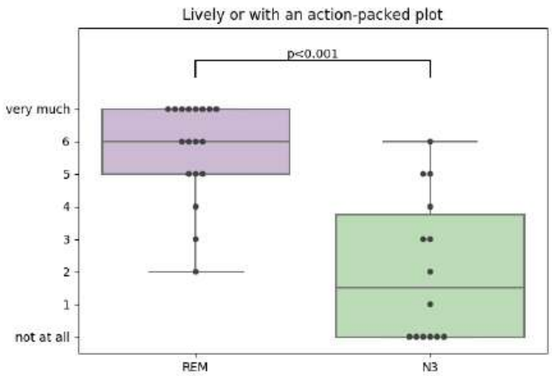


Figure 3.43: Difference between REM and N3 sleep dreams, regarding in how far they were lively or with an action-packed plot.

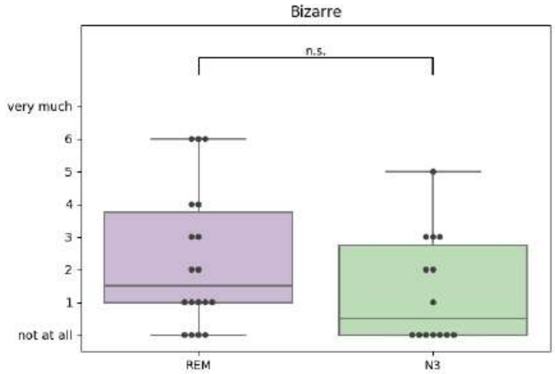


Figure 3.44: Difference between REM and N3 sleep dreams, regarding in how far they were bizarre.

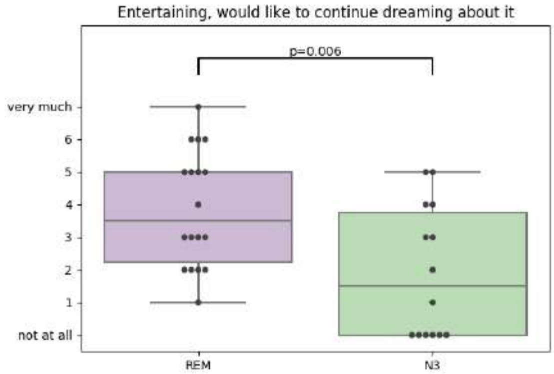


Figure 3.45: Difference between REM and N3 sleep dreams, regarding in how far they were entertaining and whether the subject wanted to continue dreaming about it.

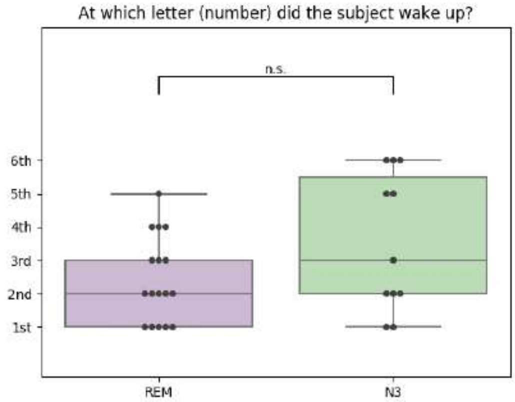


Figure 3.46: Difference between REM and N3 sleep, regarding which of the letters the subject heard first during the arousal procedure.

3.2.4.8 Sleep depth

The correlation analysis, which set the night number (1st, 2nd, 3rd) in relation to the time it took the subject to wake up, revealed that subjects slept significantly deeper during the later nights of the experiment, even though only a weak correlation could be found ($r=0.21$, $p=0.025$). Furthermore, the same type of analysis showed that acoustic stimuli presented later during a night (i. e. more towards the morning) woke up the subjects significantly faster (again only weak correlation, $r=-0.19$, $p=0.041$).

3.2.4.9 Workload of the experimenter and costs of the experiment

Before handing out the experiment boxes to the subjects, the experimenter had to extend the Traumschreiber system software slightly by adding the experiment-specific auditory arousal procedure (letters with increasing volume), which also included recording the acoustic instructions to fill out the nightly questionnaire.

When the experiment boxes were handed out to the subjects, the experimenter had to answer only some smaller organizational questions of the subjects. These were mostly about how, where and when to return the materials, and how to receive the signature for the subject hours.

During the experiment week, only one subject contacted the experimenter (via email). This was the subject with the on-off switch, which fell off before the first experimental night. A spare experiment box was offered to the subject, but she declined and chose to not continue the experiment.

All subjects returned their experiment box in time. The boxes were checked for completeness (all complete), and the sleep masks were inspected, whether any signs of wear could be detected (see above). Finally, the questionnaires and the micro-SD cards of the minicomputers were taken from each box for the data analyses.

Thus, the workload of the experimenter for actually conducting the experiment was extremely low compared to laboratory polysomnographic studies with acoustic stimulation of the subjects and sleep staging in real-time during the night.

The variable costs per night consisted of the costs of the single use sticky electrodes and the printing costs of the subject information leaflet, the consent form and the questionnaires. No money was paid to the subjects. Thus, in total less than two EUR had to be invested for each experimental night, resulting in total costs of less than 100 EUR for the whole experiment of 39 recorded nights.

4 Discussion

4.1 General discussion

The here developed Traumschreiber system offers a new methodological approach to sleep and dream research. It enables naive crowd subjects to conduct complex polysomnographic experiments at home, i. e. in their natural sleep environment. No professional human experimenter is necessary at the place of recording, which allows the researcher to focus on creating experiment ideas and on analyzing experiment results instead of spending time with repetitive tasks during the night. The Traumschreiber system instructs the subject, what to do how and when, and carries out experiment tasks, when the subject is asleep. The technology used inside the Traumschreiber system records data of good quality on eight electrophysiological channels, including EEG, EOG, EMG and ECG. Furthermore, the raw data is accessible in real-time, and real-time stimulation of the sleeping subject is possible (acoustically, visually). The Traumschreiber system is a completely open source and non-commercial technology and can be extended as needed. A detailed description, of how to create own entities of it, is supplied in this dissertation. Due to the very low costs of the system (less than 230 EUR for the complete system including a Raspberry Pi 3 Model B minicomputer, or less than 140 EUR for just the sleep mask), many Traumschreiber systems can be used even on a low budget in parallel. As a result, far more nights can be recorded by using a crowd-based approach in a much shorter time at much lower costs and with far less efforts for the researcher and the subject, than would be possible with polysomnographic sleep laboratory recordings – and at the same time going far beyond the possibilities of previous portable devices for field studies, because also complex experimental paradigms can be carried out. This enables new, large-scale and data-driven scientific explorations in the field of sleep and dream research.

The Traumschreiber system is not only far less expensive, but compares well to the previously existing sleep and dream research methods in many aspects (for further details regarding the existing technologies, please see chapter *1.3 Assessment of sleep and dreaming in the sleep laboratory* and chapter *1.4.2 Field studies using electronic tools*):

- Compared to polysomnographic in-laboratory recordings, the main advantages of the Traumschreiber system are that no laboratory is necessary, resulting in drastically reduced resource needs and workload of the experimenter. Moreover, the sleep of the subject can be studied in the natural sleep environment, avoiding potential influences of the unfamiliar setting on sleep or dreams. The visual stimulation capabilities of the high-tech sleep mask and the easy to program experiments are further advantages of the system over standard in-laboratory sleep recordings. It has to be noted, however, that the data quality (even though is it good) is worse than what can be achieved using polysomnographic sleep laboratory recordings.
- Compared to medical portable sleep recording devices, the Traumschreiber system has a lot of advantages and only a few disadvantages. Actigraphy is not able to measure EEG or EOG and thus is not able to differentiate between sleep stages, which is possible with the Traumschreiber system. Portable polysomnographic medical devices are far more expensive than the Traumschreiber system and need a trained sleep technician to apply them in the evening, in contrast to the here developed system. Moreover, to the author’s knowledge, no portable sleep experiment system exists, which allows to conduct as complex sleep experiments as is possible with the Traumschreiber system (e. g. acoustic stimulation of the sleep based on the real-time data, real-time sleep evaluation using machine learning algorithms such as deep neural networks, easy-to-program sleep experiment framework). Furthermore, the available medical portable sleep recording devices are proprietary, commercial products, limiting their extendability and, in most cases, the access to the raw data in real-time. The Traumschreiber system, on the other hand, is open-source and can be extended and modified by any interested researcher. It has to be noted, however, that the Traumschreiber system is not intended to be used for medical purposes but only designed to be used for research purposes. Moreover, there is no commercial, profit-oriented producer behind it, no salesmen and, thus, only limited technical support.
- Compared to the variety of consumer sleep recording devices, and in contrast to most of them, the Traumschreiber is validated against a polysomnographic in-laboratory device, it records more electrophysiological channels, it is cheaper and

requires no licensing fees, is open-source and allows real-time data access, is tailored to scientific sleep experiments and supports the researcher with an easy way to create automated sleep experiments, and it is ready-to-use and not only promised to be sold sometime in the nearer future. No other consumer product sleep recording device offers the combination of advantages of the Traumschreiber system. It has to be noted, however, that there is no profit-oriented company behind the Traumschreiber system offering professional user support, but only the developing researchers.

As described in chapter *1.6 Sleep and dream research and artificial intelligence*, various approach to use methods from the field of artificial intelligence for sleep research purposes exist, e. g. for using deep neural networks for sleep stage classification. For the first time, the Traumschreiber system offers an open-source hardware and software combination, into which these existing tools can be integrated – the Keras/tensorflow deep neural network framework of the Traumschreiber system can be easily filled with state-of-the-art or future neural network architectures.

Summarized, the Traumschreiber system is a new tool for sleep and dream researchers, offering a unique combination of features, which make possible new types of experiments.

4.2 Discussion of the development process and evaluation studies

4.2.1 Discussion of the development process

Starting from scratch, first the requirements were defined - based on the experiences of the author from work in two scientific sleep laboratories as well as from sleep research field studies. Furthermore, the requirements were continuously updated based on the feedback of other researchers from collaborations, conferences and workshops. It is still likely, that the features of the here presented Traumschreiber system are not optimal for every use case. This means, that other researchers would have set different priorities for development, e. g. to implement less or more recording channels, to use a better but more expensive microcontroller, to build in other types of stimulation into the sleep mask, or to set a stronger focus on developing a smartphone app. The here developed Traumschreiber system tries to optimize a complex network of interdependences between costs, quality, features, comfort, development and production time, and many more. But since the Traumschreiber system is completely open regarding both hardware and software, any researcher with enough technological background (or colleagues) can modify it according to his or her special needs.

Choosing an iteration-based approach for developing the Traumschreiber system is quite standard for high-tech devices, as everbody knows from other technical devices such as smartphones or even polysomnographic recording equipment, which have new, additional features in every new version. Overall, this approach worked fine and lead to constant improvement of the latest prototype, resulting in the final system as presented here. However, only a handful of researchers were involved in the actual development work of this complex technology, and all of them worked only part-time or as a hobby on this project. No professional electrical engineer, no professional sleep researcher with decade-long experience in the field and no professional product designer was part of the development team. Moreover, the project's budget was low, and acquiring money from independent charity organizations was a time consuming and bureaucratic act. Thus, the speed of development was rather low, and some decisions were made, which experienced industrial developers might not have made. One example for this is the decision to construct a self-designed differential EEG with operational amplifiers, instead of using ("pre-built") instrumentation amplifiers, which was neither cheaper nor better regarding the data quality in the end. Another example was to choose a cheaper factory with shorter lead times for the second PCB, which was produced and equipped in China. In the end, 9 out of 10 boards did not function properly, and the expected shorter lead time was much longer than promised. Summed up, a lot of time-consuming learning was necessary throughout the whole development process, due to starting as greenhorns. On the other hand, this probably eased innovative development research, as no "this is not possible" or "we never worked this way" stopped producing new solutions to occurring problems.

4.2.2 Evaluation studies

The Traumschreiber system was evaluated in several studies. In the first study, simultaneous sleep recordings of the Traumschreiber system and a commercial medical polysomnographic system were carried out in a sleep laboratory, and the data of both systems was compared. It was found out, that the Traumschreiber system can record data of high quality, in which the most important types of sleep graphoelements can be identified, and which lead to the classification of the same sleep stages in more than 90% of the compared epochs of both systems. In the second study, a rather simple type of sleep experiment was conducted by naive crowd subjects using a prototype of the system, demonstrating that this approach is in principle feasible. In the third study, it was measured, what resources are necessary for producing the Traumschreiber system, and the work steps were described in detail, so that any interested researcher can produce an arbitrary number of Traumschreiber systems. This description also functions as a tutorial. The last evaluation study showed, that many Traumschreiber systems can be used in parallel by multiple naive crowd subjects with very low organizing demands for the scientist. Moreover, it was shown that the built-in machine learning based algorithms are able to steer complex experimental paradigms, which also require real-time analysis of the subject's sleep (even though improvements of the internal sleep staging algorithm are needed).

Some limitations of the general evaluation process need to be addressed, before each of the four studies is discussed. First, the evaluation studies were conducted by the developers of the Traumschreiber system and their students. Even though all studies were carried out in a highly objective way and following good scientific practices, independent evaluations of the Traumschreiber system should be conducted.

Second, another weakness of these studies is the selection of the subjects. Only healthy young adults participated in the three studies, which evaluated the Traumschreiber system in independent subjects. Thus, it remains unclear, how well the Traumschreiber system is suited for measuring patients with sleep disorders or other illnesses, for who sleep staging might be harder, or older subjects, who might be more insecure with using this new technology.

Another limitation regarding the evaluation studies is the usage of different versions of the Traumschreiber system. Of course it would have been a better idea to conduct these studies all with the exact same device after finishing development, instead of using a prototype for two studies. However, due to practical reasons this was not possible (the project's time frame and the iterative development process). And since very far developed prototypes were used, which differed not significantly regarding the relevant technical specifications, this procedure was viewed as an acceptable solution.

4.2.2.1 First study: Sleep laboratory validation

This study simultaneously measured sleep of six subjects with both the Traumschreiber system and the ALICE system, a commercial polysomnographic system, which is used in clinical and research sleep laboratories worldwide. The comparison of 189 epochs (30 second data bins) of both systems demonstrated, that both systems record very similar data, and that the Traumschreiber system can, in principle, thus be used for sleep experiments. All sleep graphoelements (e. g. slow waves, sleep spindles, K-complexes, REMs, heart beat) could be detected in both systems in nearly all corresponding epochs, e. g. if one system showed slow waves, the other systems showed slow waves, too. The only exceptions to this were central-occipital sleep spindles and alpha EEG activity, which were mostly not detectable in the Traumschreiber system due to the frontal placement of the electrodes. However, if one changes the suggested recording layout and collects EEG data with occipital electrodes (e. g. cup electrodes), these graphoelements should be measurable as well. The downside of this idea is, that inexperienced subjects will find it very hard to impossible to fixate the cup electrodes on the back of their head.

Manually sleep staging the 189 epochs showed, that in 92.6% of the cases, the same sleep stage would be assigned. This demonstrates, that the data of the Traumschreiber system is well suited for professional sleep experiments, which want to analyze or compare only selected sleep stages. However, viewing only single 30-second epochs for classifying sleep stages, instead of scoring the complete nights, led to the effect that sometimes it was impossible two differentiate between N1, N2, REM sleep or wakefulness, since knowledge about the sleep stage of the previous epoch would have been necessary. Thus, an ambivalent sleep category had to be introduced for these cases. This does not necessarily mean, that sleep staging these epochs is not possible, however, it remains unclear, how well the Traumschreiber system can be used in these complicated cases.

Another limitation of this experiment is, that only seven subjects were recorded, with one of the recordings being excluded afterwards. Thus, sleep characteristics, which are rare to find in many subjects, might have been missed (e. g. sawtooth waves in the sleep EEG of REM sleep).

The placement of the electrodes in this experiment has to be discussed as well. As can be seen in figure 2.14, the electrodes of both systems were placed as close to each other as possible. However, because of the size of the single use sticky electrodes, about two to three centimeters of space remained between the electrodes of the two systems, which might be an explanation for the slight differences in the recordings of both systems. Another reason for the slight differences in the shape of both signals might be special hardware or software filters of the ALICE system, which are not known since the system is proprietary (not open).

4.2.2.2 Second study: Home-based prototype study

In this study, the Traumschreiber system was used for the first time by naive crowd subjects. The results of this experiment show, that home-based autonomous sleep experiment including acoustic stimulation of the sleep of subjects, who are instructed by the system how to set up the polysomnographic measurements, are feasible, and that the subjects find it rather easy to use the Traumschreiber system. Moreover, this experiment identified several points, how the system needed to be improved. Most importantly, an automatically re-establishing bluetooth connection was necessary. The experiment shows further, that the data quality of the recordings measured by the Traumschreiber system was good, if the bluetooth connection did not disconnect during the night.

The sleep quality was significantly better during the control night, in which the sleep mask was not worn, than in the experimental night with the Traumschreiber system. This was the case for the experimental subgroup, which received acoustic stimuli every ten minutes throughout the whole night, as well as for the subgroup, which was not stimulated acoustically. The feeling of recovery in the morning, however, was not significantly different between the control night and the night with wearing the Traumschreiber sleep mask for the subgroup without acoustic stimulation (but for the subgroup, which was stimulated, it was). Moreover, the majority of the subjects reported, that the sleep mask was not or only slightly disturbing and was at least okay to wear, if not even comfortable. Taken together, these results suggest, that the Traumschreiber system leads to slightly impaired sleep, even though wearing it was comfortable or at least okay for the subjects.

Not surprisingly, acoustic stimuli, which lead to arousals, were significantly louder than the cues, which did not provoke an arousal. Even though this outcome was to be expected, this can be viewed as a further sanity check, that the Traumschreiber system conducted its task successfully. Another insight of finding such an effect is, that even though differences between the subjects and their sleeping environment existed and a direct comparison between them was not sensible, on a group level significant effects could be found. Especially for much larger crowd experiments with many more subjects, this effect is a strength of the Traumschreiber system, as it is inexpensive and can be recorded simultaneously. This means, that for example differences regarding the exact distance between the speakers and the bed, the direction of the speakers, the individual preferred body positions, and other circumstances such as the type of pillow or the general loudness of the room, can be neglected, if enough subjects are recorded.

A limitation of this study is, that only one night was recorded for each subject with the prototype of the Traumschreiber system, and one control night without the system. No adaptation or practice night was conducted. As described in the results (how to improve the system), this limits the study's validity regarding sleep quality, ease of use, and also data quality, since the subjects might have got used to wearing the sleep mask. Note, however, that in the fourth study, these questions were analyzed in greater detail.

4.2.2.3 Third study: Preparing polysomnographic crowd experiments

This study analyzed, how quickly and at which costs a Traumschreiber system can be produced, and described each individual work step. The description of the work steps can be viewed as a tutorial for other researchers, who want to use this new, open technology themselves. Overall, the production time for one Traumschreiber system was found to vary strongly depending on the experience of the manufacturer, and can be assumed to lie between 170 minutes for inexperienced workers to 60 minutes for people experienced in producing the system, especially if larger quantities are produced and several Traumschreiber systems can be produced in parallel. Additionally, several weeks of time have to be kept in mind for waiting time, until the ordered components arrive. One limitation of this study is that the experimenter, who conducted the work steps, was part of the Traumschreiber system development team (the author of this dissertation). Thus, it can only be estimated, how long each work step takes for completely inexperienced experimenters. On the other hand, having a tutorial at hand like the one presented in this dissertation probably increases work speed substantially, as best practices are available. Moreover, it can be criticized, that the fixed number of Traumschreiber systems, which were produced during this study (20 pieces), and the total number of producers (one person) limits the generalizability of this study further.

The production costs of the Traumschreiber system are very low, compared to commercial devices, which, moreover, do not even offer the same functionality (e. g. open hardware/software, raw data access in real-time, artificial intelligence controlled experiment conduction). 20 of these low-budget systems cost 222 EUR each (as of March 2017). If the Raspberry Pi experiment station is replaced by a smartphone app and mobile phones of the subjects are used, the production costs drop by further 34%. However, the existing Android app protoype needs substantial further development, before it can be used in a similar way as the Raspberry Pi experiment station. Another big cost factor are the electrode cables, which cost around 40 EUR, and which could be imported from China for around a third of the price, as a sample order showed. One has to keep in mind, however, that all these costs depend on several factors, ranging from the total order quantity to the currency exchange rate, and can thus only serve as estimators. Furthermore, the production costs include only those costs for the materials. Depending on the situation of the researcher producing one or more Traumschreiber systems for own experiments, additional costs for paying production assistants, premises and so on might occur.

Taken together, this study gives a detailed description of the work steps necessary to produce arbitrarily many Traumschreiber systems. The production times and costs should be seen as an estimate, giving the researcher a first impression of the resources needed. They also enable the researcher to compare the production time for a Traumschreiber system to the time usually spent for recording a single night in the sleep laboratory: During the time needed for recording one subject in the laboratory, about eight Traumschreiber systems can be produced.

4.2.2.4 Fourth study: A crowd-based polysomnographic experiment

The most complicated experimental paradigm was used in the last of the four evaluation studies. More than ten subjects received one Traumschreiber system each and recorded three nights each, demonstrating the system's suitability for parallel recording. Despite obtaining 39 polysomnographic datasets within one week, the efforts and the costs for the experimenter stayed very low, demonstrating the scalability aspect of crowd-based sleep experiments. Moreover, several other evaluations were conducted.

First, in general, the data quality was found to be very good in nearly all nights, if the data transmission worked well and the sleep masks' power supply was stable. However, this was not the case in most of the first nights of each subject. Loose on-off switches in three of the fourteen handed out sleep masks as well as not fully charged batteries (in the first night) were major sources for nights with unusable data. It might also be the case, that some subjects intentionally or accidentally switched off the sleep mask throughout the first night. In general, the recording quality of the second and third nights were much better than during the first night. This shows, that an adaptation and practice night is helpful.

Ten of the eleven subjects with intact sleep masks could deliver a fully functional polysomnographic recording for at least one of the second or third night (N=22). Averaged over all subjects with intact sleep masks for the second and third night, 86% of the recordings showed a good EEG signal, 86% a good EMG signal, 82% a good EOG signal, and 68% a good ECG signal. Nevertheless, a quite high fraction of recordings with not complete polysomnography (i. e. EEG, EOG and EMG) after the practice and adaptation night remains (6 of 22 nights), despite the number of backup channels. As a result, researchers need to be aware in future studies, that such autonomous recordings conducted by inexperienced subjects lead to a quite substantial amount of dropout nights. This is might not be very problematic for most experiments, as recording a few further nights is inexpensive and results only in a limited amount of extra work and costs for the experimenter, but might be a disadvantage, if a specific night of a specific subject is of importance.

The sleep quality and the feeling of recovery in the morning as measured using the Schlaffragebogen A questionnaire were found to be similar in all three experimental nights. However, when the subjects were asked directly during the later nights, whether they slept better or worse than in previous experiment nights, the average answer was "slightly better". Moreover, the subjects were significantly harder to wake up during later nights of the experiment by the Traumschreiber system's acoustic stimulations. Cautiously interpreting these contradicting results, it can be concluded that the subjects' sleep was either equally good or slightly better during the later nights, but definitely not worse. The mixed findings regarding the sleep quality differences between the three experimental nights should be investigated further in future studies, which might also analyze this effect for a longer time period than three nights. Furthermore, it was found out that the sleep quality was significantly worse during the here recorded nights than in the questionnaire control group, data of which was provided by the questionnaire authors. This is not surprising, though, because they subjects were woken up several times during each night by the Traumschreiber system on purpose.

The crowd subjects reported, that the Traumschreiber system is easy to use. The instructions video, which showed how to put on the electrodes and the sleep mask, was very helpful. In all nights, no major technical problems (apart from the loose on-off switches in three sleep masks) were reported. The subjects found the sleep mask neither comfortable nor uncomfortable, but slightly disturbing for falling asleep. Except one broken ECG cable, no signs of wear could be detected. This demonstrates, together with the overall good data quality, that even sleep recording inexperienced subjects can successfully measure their sleep on their own, if they are instructed as done by the Traumschreiber system, and return these data a few days later to the experimenter.

The next evaluation showed, that the Traumschreiber system performed a real-time sleep stage classification in all recorded nights using a deep neural network approach. The system also adjusted the experiment timings accordingly, and took subject-individual differences into consideration. However, the prediction quality of the simplistic sleep staging algorithm was by far not sufficient and needs further improvement. The bad sleep stage prediction quality was to be expected, since the machine learning algorithm was trained on only a few nights of data and was constructed in a very simple form. Nevertheless, the underlying technology (deep neural networks using Keras/tensorflow) has been shown to be able to solve very complicated tasks (see chapter *1.5.2 Machine learning and deep neural networks*), including image and video classification, and including tasks, for which several temporal timescales have to be considered. If such an advanced deep neural network is trained, it can be run on the Traumschreiber system without any modifications, promising excellent sleep staging quality at least for healthy human sleep.

An exemplary dream research experiment was conducted in this study as well, and showed, that one can come to similar results by using the Traumschreiber system as with in-laboratory sleep research. In this experiment, differences between 18 REM and 14 N3 dreams were analyzed. REM dreams were found to be significantly less static, but more action-packed and entertaining, fitting results from previous dream research (Nir & Tononi, 2010). Even though REM dreams were rated as being more bizarre, this effect was not statistically significant, probably due to the low number of dreams being analyzed. The same is true for the time needed to wake the subject up: There might be a difference between the two sleep stages, but this was not found to be statistically significant.

Finally, taking all the before mentioned points together, the experiment showed, that the Traumschreiber system

- enables naive crowd subjects to conduct complex, sleep-stage dependent polysomnographic recordings including stimulation in a home-setting, and
- that is can take over the experiment lead during the time the subject is asleep.

The general strength of this crowd-based, large-scale, data and machine learning driven approach is, that despite a possibly higher dropout rate and more nights with bad data quality, the sheer amount of nights, which can be recorded without much effort, leads to a high number of successful recordings, and possibly new scientific insights about sleep and dreaming.

4.3 Future research directions and prospects

4.3.1 Future research directions

4.3.1.1 Further developing the Traumschreiber system

The main goal of this dissertation could be reached: The Traumschreiber system enables naive crowd subjects to conduct complex interactive polysomnographic sleep experiments at home, without the need for a professional researcher to be at the place of recording. Furthermore, it serves as a basis for further open source developments and even more elaborated experiments. Nevertheless, the system can be improved in nearly every aspect. This is due to the complexity of this project, limiting the available time for all the different subtasks – which ranged from identifying modern sleep and dream research's limitations, to electrical engineering of a whole polysomnographic system including EEG, to costs optimization including outsourcing to Chinese factories, to increasing the usability of the system and comfort of the subject, to data quality analyses, to public relation management, to obtaining finances for the project, to conducting the actual sleep studies with the new system, to training and supervising research assistants, to actually managing all this, setting the right priorities and keeping an eye on the overall timely constraints.

Despite the great support of several co-workers with the subtask for this project, quite a number of tasks and improvements remain open. These are:

- Better sleep staging and pattern recognition algorithms: Even though simplistic sleep staging and pattern recognition algorithms based on machine learning are implemented into the Traumschreiber system, the goal of this dissertation was not to optimize their performance. Thus, the existing underlying technology based on Keras/tensorflow deep neural networks can be used much more effectively. As was described in chapter *1.6 Sleep and dream research and artificial intelligence*, numerous approaches for automatic sleep staging and pattern recognition exist. A promising idea might be to combine the excellent pattern detection abilities of convolutional neural networks (LeCun, Bottou, Bengio, & Haffner, 1998) with the multi timescale properties of long-short-term-memory cells (LSTMs, (Hochreiter & Schmidhuber, 1997)). Also, the preprocessing and feature extraction of the data can be improved, for example in a similar way as described in (Popovic, Khoo, & Westbrook, 2014). This way, both long term patterns in the data like sleep cycles as well as short term pattern such as the current spindle activity would be used.
- Conducting crowd experiments: The way, how the crowd experiments are conducted, could be elaborated on further. For example, by implementing a better subject interaction (adjusting volume, repeating instructions) or by showing the instruction video directly on a screen of the experiment station instead of sending the subject an internet link to the video. Moreover, the questionnaires could be conducted in a digital form, and the experiment station could control, that the subject fills them out at the correct time point (e. g. by keeping the subject awake until the questionnaire is completed during the night). Furthermore, the Traumschreiber system could check, that the battery of the sleep mask is charged before the recording starts and that the data quality is sufficient (good electrode contact).
- Hardware improvements: Regarding the hardware, several further developments should be thought of. More data could be transferred over the BLE connection by using compression methods, leading to a higher sampling rate. Furthermore, an impedance measurement could be implemented. Lastly, a different battery holder as well as a more stable on-off switch could be used.

- Data encryption: A dedicated encryption chip is already placed on the PCB, but has not been used so far. The idea is to encrypt the recorded data directly in the sleep mask, send it to any real-time data analysis station (possibly even over the internet, as shown in (Kayyali et al., 2008)), and to decrypt the data there. A similar approach for encrypting sensor data in general was already proposed by (Bruyneel, Van den Broecke, Libert, & Ninane, 2013; Healy, Newe, & Lewis, 2008).
- An Android smartphone app to be used instead of a Raspberry Pi: A demo app was programmed and receives BLE data from a dummy device (see chapter *3.1.2.1.2 Experiment station*). The BLE service can be run in background without stopping data transmission. Furthermore, some introductory slides are available to the user and can be modified. The next steps are: modifying the code so that it receives BLE data from the Traumschreiber sleep mask, implementing a basic experiment station functionality like storing the data on microSD card and plotting the real-time data, and eventually adding real-time data analysis features (the tensorflow neural networks are in principle usable on Android as well (Abadi, Agarwal, et al., 2016)). For future versions, one could think of using the Android app for forwarding the (encrypted) data over the internet to a server and conducting the data storage and real-time data analyses there, sending back only the results and commands to the smartphone, similar to approaches in telemedicine (Craig & Patterson, 2005).
- Reusable electrodes: During the textile development of the Traumschreiber sleep mask, several ideas how to replace the single use sticky electrodes by electrically conductive fabric or yarn have been tried out, as they have been successfully used for example by the Zeo consumer product (Griessenberger et al., 2013; Shambroom et al., 2012). This included stitching reusable electrodes into the sleep mask, using conductive fabric as electrodes, and replacing the electrode cables by conductive yarn. It became clear, however, that the signal quality of the self-stitched electrodes was not good enough yet. Further research should investigate this idea in more detail, as this could make the system easier to use and more comfortable.
- Safety certificates: Depending on the actual usage, specific safety measurements and certificates might be necessary, for example regarding electromagnetic compliance. Using the Traumschreiber system for medical purposes requires additional safety checks and certificates.

- The look and feel of both the Traumschreiber sleep mask as well as the Traumschreiber experiment station could be improved further.

4.3.1.2 Suggestions for further studies with the Traumschreiber system

On the one hand, studies should be conducted, which evaluate the Traumschreiber further:

- A study, which compares the sleep quality of nights with the Traumschreiber system, with in-laboratory sleep and with undisturbed sleep, all in the same subjects, should be conducted. This way, it could be assessed, how large the effect of the Traumschreiber system is on the sleep of the subjects. Such an extended validation of the Traumschreiber system should be conducted in independent sleep laboratories. Besides comparing the sleep graphoelements and manually scored sleep stages, the comparison could also include an evaluation of the (at that time possibly more advanced) automatic sleep staging capabilities of the system, and include other sleep characteristics like the total sleep time, sleep efficacy, or sleep latency (compare (Popovic et al., 2014)).
- Moreover, a large-scale study with at least 100 subjects, which runs over a longer period of time (at least ten nights per subject), should be conducted. In a pessimistic view, if only half of the recorded nights produce polysomnographic data of good quality, this would result in 500 successfully recorded nights. These could be scored using the quick scoring method of the Traumschreiber system, and then be used as training data for improving the automatic sleep scoring capabilities of the Traumschreiber system – for sleep staging, but also for other pattern detection methods. Moreover, the durability of the material could be evaluated in such a long-term study. Assuming that the subjects can be recruited without any difficulties and 20 Traumschreiber systems are at hand, these 1000 nights could be recorded (in theory) in less than two months, with 3 to 6 months being a more realistic estimate.

On the other hand, the Traumschreiber system makes possible several new experiments. A few of these ideas are listed here in order to demonstrate their variety:

- First, in the field of memory consolidation, complex sleep laboratory studies could be replicated by a subject crowd in the field, building a bridge from the theoretical in-laboratory research into the practice. For example, a study (Ngo, Martinetz, Born,

& Mölle, 2013) showed, that a specific type of auditory stimulation during slow wave sleep increases the memory performance (vocabulary) by around 70%. The Traumschreiber system could be used to replicate this finding in the field, by implementing this type of real-time pattern detection and real-time stimulation into it. Moreover, by using the built-in LEDs of the Traumschreiber sleep mask, the same experiment could be conducted with a different modality and the results could be compared. Furthermore, it could be investigated, how easily this technology can be transferred to the masses and to a crowd outside of science.

- Second, the role, which sleep plays for gaining insights into current problems, could be investigated further using the Traumschreiber system. A study (Wagner, Gais, Haider, Verleger, & Born, 2004) suggests, that sleep plays an important role for restructuring new memory representations and facilitates the extraction of explicit knowledge. However, this study could not be replicated so far. Recording a larger amount of subjects with the Traumschreiber system and conducting this type of experiment in an automated way could lead to new insights about insights inspired by sleep. Moreover, the experiment could be extended by investigating a possible effect of auditory or visual cueing, as suggested in (Appel et al., 2016).
- Third, the effect of auditory or visual stimulation on the dream content could be investigated. A number of studies in this direction has already been carried out, e. g. (Dement & Wolpert, 1958). Since the costs of such studies using the Traumschreiber system are very low, extending these experiments and including far more types of acoustic or visual stimulation in far more subjects might lead to new insights about dreaming in general, but also about how dreams can be altered from the waking world, possibly opening new research perspectives for nightmare therapy.
- Fourth, if the Traumschreiber system is used for medical purposes, several medical assessments and therapies could be further developed and conducted by naive subjects or patients in the field. For example, the multiple sleep latency test (MSLT, (Carskadon, 1986)) could be conducted in such a way (even though it would need adjustment to the recording environment), or new depression therapies could be developed by inhibiting sleep or specific sleep stages in a specific way (Dallaspezia & Benedetti, 2015). Moreover, if sleep physicians are provided with many nights of

a patient's sleep in the natural environment, instead of just one or two in-laboratory nights, better diagnoses could become possible.

- Fifth, lucid dreaming experiments could benefit greatly from the Traumschreiber device, since lucid dreams are very hard to record in the sleep laboratory (compare chapter *1.3 Assessment of sleep and dreaming in the sleep laboratory*). Thus, by using the Traumschreiber system, elaborated lucid dreaming techniques such as the communication between the lucid dreamer and the waking world, in which eye movements and acoustic or visual stimuli are used (Appel, 2013), could be practiced by the subject at home for several nights, before recording them in the laboratory.
- Lots of further research ideas can be thought of, e. g. in the field of chronobiology and jet-lag (e. g. applying light stimuli (Geerdink, Walbeek, Beersma, Hommes, & Gordijn, 2016)). More ideas are described in (Braumann, 2016).

4.3.2 Future prospects

Even though the Traumschreiber system is designed for scientific sleep and dream experiments, it might also be used for non-research purposes. One could imagine, that one day many people would have their own personal sleep assistant (Daskalova et al., 2016), who knows the individual sleep characteristics, knows how to stimulate the sleep best in order to increase the sleep quality, and maybe even knows how to enrich the dreams of the user. Devices based on the Traumschreiber system could be used for this.

Due to its very low costs, the Traumschreiber system is an ideal candidate for using it for educational purposes. This could benefit students in disciplines such as electric engineering, neuroscience, medical science or computer science. The great advantage of the Traumschreiber system is, that the students would have a direct connection to the recorded data (their own EEG / ECG / EOG / EMG activity), making the topic to be learned more interesting to them.

Also, using the Traumschreiber system for EEG, ECG or eye tracking studies during wakefulness might be an interesting option. Moreover, the Traumschreiber system could become an inexpensive and easy to use tool also for researchers from other disciplines, who are unfamiliar with the sleep laboratory environment, but who would like to analyze the subjects' sleep as a side note.

As the Traumschreiber system is open source, a crowd of developers, researchers and consumer users might grow, and the creativity and work force of these people could be bundled.

Many more use cases can be thought of. The most important point regarding the future of the Traumschreiber system is, however, that it is actually used. It does not matter, for which purpose: for sleep and dream studies, for the personal fun of single developers and users, for education, or for completely different purposes.

5 References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., … Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv preprint arXiv:1603.04467.*

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Kudlur, M. (2016). TensorFlow: A System for Large-Scale Machine Learning. In *OSDI,* 16, 265-283.

Agnew, H. W., Webb, W. B., & Williams, R. L. (1966). The first night effect: An EEG study of sleep. *Psychophysiology*, 2(3), 263–266.

American Academy of Sleep Medicine (2014). International Classification of Sleep Disorders: Diagnostic and Coding Manual. 3rd edition (ICSD-3). *AASM Resource Library*, 281(9).

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. P. (2003). The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, 26(3), 342–92.

Appel, K. (2013). Communication with a Sleeping Person. *Unpublished Master's thesis,* Osnabrueck University.

Appel, K., De Vries, E., Hildebrandt, K., Immel, A. S., Ivanovs, M., Kar, B., … Pipa, G. (2016). Effects of Sleep and Cueing on Insight. *Unpublished study project report,* Osnabrueck University.

Appel, K., Pipa, G., & Dresler, M. (2017). Investigating consciousness in the sleep laboratory - an interdisciplinary perspective on lucid dreaming. *Interdisciplinary Science Reviews*. 1-16.

Babaeizadeh, S., White, D. P., Pittman, S. D., & Zhou, S. H. (2010). Automatic detection and quantification of sleep apnea using heart rate variability. *Journal of Electrocardiology* 43(6), 535–541.

Braumann, S. (2016). Sensory stimulation during sleep: an overview of past experiments and promising future directions. *Unpublished Bachelor's thesis,* Osnabrueck University.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Brown, P., Cocke, J., Pietra, S. Della, Pietra, V. Della, Jelinek, F., Mercer, R., & Roossin, P. (1988). A statistical approach to language translation. In *Proceedings of the 12th conference on Computational linguistics,* 1, 71–76.

Bruyneel, M., & Ninane, V. (2014). Unattended home-based polysomnography for sleep disordered breathing: Current concepts and perspectives. *Sleep Medicine Reviews*, 18(4), 341–347.

Bruyneel, M., Van den Broecke, S., Libert, W., & Ninane, V. (2013). Real-time attended home-polysomnography with telematic data transmission. *International Journal of Medical Informatics*, 82(8), 696–701.

Bsoul, M., Minn, H., Nourani, M., Gupta, G., & Tamil, L. (2010). Real-time sleep quality assessment using single-lead ECG and multi-stage SVM classifier. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10* (pp. 1178–1181).

Buysse, D., Reynolds, C., Monk, T., Berman, S., & Kupfer, D. (1989). The Pittsburgh Sleep Quality Index (PSQI): A new instrument for psychiatric practice and research. *Psychiatry Research*, 28, 193–213.

Callini, G., Essig, S. M., Heher, D. M., & Young, L. R. (2000). Effectiveness of an expert system for astronaut assistance on a sleep experiment. *Aviation, Space, and Environmental Medicine*, 71(10), 1023–32.

Campbell, A. J., & Neill, A. M. (2011). Home set-up polysomnography in the assessment of suspected obstructive sleep apnea. *Journal of Sleep Research*, 20(II), 207–213.

Carskadon, M. A. (1986). Guidelines for the Multiple Sleep Latency Test (MSLT): A Standard Measure of Sleepiness. *Sleep*, 9(4), 519–524.

Chollet, F. (2015). Keras. Retrieved from https://github.com/fchollet/keras

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1). 155–159.

Craig, J., & Patterson, V. (2005). Introduction to the practice of telemedicine. *Journal of Telemedicine and Telecare*, 11(1), 3–9.

Crisler, S., Morrissey, M. J., Anch, A. M., & Barnett, D. W. (2008). Sleep-stage scoring in the rat using a support vector machine. *Journal of Neuroscience Methods*, 168(2), 524–534.

Dallaspezia, S., & Benedetti, F. (2015). Sleep deprivation therapy for depression. *Current Topics in Behavioral Neurosciences*, 25, 483–502.

Darrach, B. (1970). Meet Shaky, the first electronic person. *Life*. 56–68.

Daskalova, N., Metaxa-Kakavouli, D., Tran, A., Nugent, N., Boergers, J., McGeary, J., & Huang, J. (2016). SleepCoacher: A Personalized Automated Self-Experimentation System for Sleep Recommendations. *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 347–358.

Davis, H., Davis, P. A., Loomis, A. L., Hervey, E. N., & Hobart, G. (1939). Electrical reactions of the human brain to auditory stimulation during sleep. *Journal of Neurophysiology*, 2, 500–514.

Dement, W., & Wolpert, E. A. (1958). The relation of eye movements, body motility, and external stimuli to dream content. *Journal of Experimental Psychology*, 55(6), 543–553.

Diekelmann, S. (2014). Sleep for cognitive enhancement. *Frontiers in Systems Neuroscience*, 8, 46.

Federal Statistical Office Germany (Destatis) (2017). Cost of illness: Germany, years, disease diagnoses (ICD-10). Evas number 23631-0001.

Fell, J., Röschke, J., Mann, K., & Schäffner, C. (1996). Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures. *Electroencephalography and Clinical Neurophysiology*, 98(5), 401–410.

Fisher, R. A. (1937). *The design of experiments*. Edinburgh; London: Oliver And Boyd.

Gais, S., Mölle, M., Helms, K., & Born, J. (2002). Learning-dependent increases in sleep spindle density. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 22(15), 6830–6834.

Geerdink, M., Walbeek, T. J., Beersma, D. G. M., Hommes, V., & Gordijn, M. C. M. (2016). Short Blue Light Pulses (30 Min) in the Morning Support a Sleep-Advancing Protocol in a Home Setting. *Journal of Biological Rhythms*, 31(5), 483–497.

Ghegan, M. D., Angelos, P. C., Stonebraker, A. C., & Gillespie, M. B. (2006). Laboratory versus Portable Sleep Studies: A Meta-Analysis. *The Laryngoscope*, 116(6), 859–864.

Goble, G. (2012). Top 10 bad tech predictions. Retrieved September 10, 2017, from https://www.digitaltrends.com/features/top-10-bad-tech-predictions/5/

Goertzel, B. (2014). Artificial General Intelligence : Concept , State of the Art , and Future Prospects. *Journal of Artificial General Intelligence*, 5(1), 1–48.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1). Cambridge: MIT press.

Goril, S., & Shapiro, C. M. (2013). Rating Scales to Assess Sleep Disorders. In *Encyclopedia of Sleep* (pp. 14–19). Elsevier.

Görtelmeyer, R. (1986). SF-A und SF-B. Schlaffragebogen A und B, Selbstbeurteilungsskala (S). Internationale Skalen für die Psychiatrie. Weinheim: Beltz.

Griessenberger, H., Heib, D. P. J., Kunz, A. B., Hoedlmoser, K., & Schabus, M. (2013). Assessment of a wireless headband for automatic sleep scoring. *Sleep and Breathing*, 17(2), 747–752.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE International Conference on Computer Vision*, 1026–1034.

Healy, M., Newe, T., & Lewis, E. (2008). Analysis of hardware encryption versus software encryption on wireless sensor network motes. In *Lecture Notes in Electrical Engineering*, 20, 3–14.

Hebb, D. O. (1949). The Organization of Behavior. New York: Wiley.

Hinton, G. E., Dayan, P., Frey, B. J., & Neal, R. M. (1995). The “wake-sleep” algorithm for unsupervised neural networks. *Science*, 268(5214), 1158–1161.

Hobson, J. A., Pace-Schott, E. F., & Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *The Behavioral and Brain Sciences*, 23(6), 793–842.

Hochreiter, S., & Schmidhuber, J. J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1–32.

Horne, J. a., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int J Chronobiol*. 4(2), 97–110.

Hsu, Y. L., Yang, Y. T., Wang, J. S., & Hsu, C. Y. (2013). Automatic sleep stage recurrent neural classifier using energy features of EEG signals. *Neurocomputing*, 104, 105–114.

Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. (2007). The AASM Manual for the Scoring of Sleep and Associates Events: Rules, Terminology and Technical Specifications. (Rochester).

Jäkel, M. (2017). Developing and designing a mediator smartphone application for portable sleep EEG recording and biofeedback with focus on the end user experience. *Unpublished Bachelor's thesis,* Osnabrueck University.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep, 14(6), 540–545.*

Joseph Weizenbaum. (1966). ELIZA - A Computer Program For the Study of Natural Language Communication Between Man And Machine. *Communications of the ACM*, 9, 36–45.

Kayikcioglu, T., Maleki, M., & Eroglu, K. (2015). Fast and accurate PLS-based classification of EEG sleep using single channel data. *Expert Systems with Applications*, 42(21), 7825–7830.

Kayyali, H., Weimer, S., Frederick, C., Martin, C., Basa, D., Juguilon, J., & Jugilioni, F. (2008). Remotely attended home monitoring of sleep disorders. *Telemedicine Journal and E-Health*, 14(4), 371–374.

Kern, S., Appel, K., Schredl, M., & Pipa, G. (2017). No effect of alpha-GPC on lucid dream induction or dream content. *Somnologie*, 21(3), 180-186.

Kim, B. Y., & Park, K. S. (2000). Automatic sleep stage scoring system using genetic algorithms and neural network. In *Proceedings of the 22nd Annual Engineering in Medicine and Biology Society International Conference,* 2, 849–850.

Koley, B., & Dey, D. (2012). An ensemble system for automatic sleep stage classification using single channel EEG signal. *Computers in Biology and Medicine*, 42(12), 1186–1195.

Kolla, B. P., Mansukhani, S., & Mansukhani, M. P. (2016). Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert Review of Medical Devices*, 13(5), 497–506.

Kovalev, V., Kalinovsky, A., & Kovalev, S. (2016). Deep Learning with Theano, Torch, Caffe, TensorFlow, and Deeplearning4J : Which One Is the Best in Speed and Accuracy?, (October), 99–103. Retrieved from http://imlab.grid.by/

Kushida, C. A., Littner, M. R., Morgenthaler, T., Alessi, C. A., Bailey, D., Coleman, J., … Wise, M. (2005). Practice Parameters for the Indications for PSG—AASM Practice Parameters for the Indications for Polysomnography and Related Procedures: An Update for 2005. *Sleep*, 28(4), 499–521.

LaBerge, S., Levitan, L., Rich, R., & Dement, W. C. (1988). Induction of lucid dreaming by light stimulation during REM sleep. *Sleep Research*, 17, 104.

Lajnef, T., Chaibi, S., Ruby, P., Aguera, P. E., Eichenlaub, J. B., Samet, M., … Jerbi, K. (2015). Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines. *Journal of Neuroscience Methods*, 250, 94–105.

Laubisch, F. (2016). Extracting eye positions from electrooculographic recordings using a machine learning approach. *Unpublished Bachelor's thesis,* Osnabrueck University.

LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2323.

Lee, H., Grosse, R., Ranganath, R., & Ng, A. Y. (2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, 1–8.

Lichstein, K. L., Stone, K. C., Donaldson, J., Nau, S. D., Soeffing, J. P., Murray, D., … Aguillard, R. N. (2006). Actigraphy validation with insomnia. *Sleep*, *29*(2), 232–9.

Lighthill, J. (1973). Artificial Intelligence: A General Survey. *Science Research Council*.

Lin, H. W., Tegmark, M., & Rolnick, D. (2017). Why Does Deep and Cheap Learning Work So Well? *Journal of Statistical Physics*, 168(6), 1223–1247.

Lindsay, R. K., Buchanan, B. G., Feigenbaum, E. A., & Lederberg, J. (1993). DENDRAL: A case study of the first expert system for scientific hypothesis formation. *Artificial Intelligence*, 61(2), 209–261.

Luyster, F. S., Strollo, P. J., Zee, P. C., & Walsh, J. K. (2012). Sleep: A Health Imperative. *Sleep*, 35(6).

Mandt, L. S. (2017). Conducting a Polysomnographic, Automated Auditory Stimulation Sleep Experiment Using a Portable High-Tech Sleep Mask. *Unpublished Bachelor's thesis,* Osnabrueck University.

Marino, M., Li, Y., Rueschman, M. N., Winkelman, J. W., Ellenbogen, J. M., Solet, J. M., … Buxton, O. M. (2013). Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*, 36(11), 1747–55.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence. In *AI Magazine (2006)*, 27(4), 12–14.

McCorduck, P. (2004). *Machines who think*. A K Peters, Ltd.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115–133.

Morgenthaler, T., Alessi, C., Friedman, L., Owens, J., Kapur, V., Boehlecke, B., … American Academy of Sleep Medicine (2007). Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. *Sleep*, 30(4), 519–29.

Myllymaa, S., Muraja-Murro, A., Westeren-Punnonen, S., Hukkanen, T., Lappalainen, R., Mervaala, E., … Myllymaa, K. (2016). Assessment of the suitability of using a forehead EEG electrode set and chin EMG electrodes for sleep staging in polysomnography. *Journal of Sleep Research*, 25(6), 636–645.

Ngo, H. V. V, Martinetz, T., Born, J., & Mölle, M. (2013). Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. *Neuron*, 78(3), 545–553.

Nienhaus, F. (2017). A comparative study of a hight-tech sleep mask and a polysomnographic sleep laboratory system. *Unpublished Bachelor's thesis,* Osnabrueck University.

Nilsson, N. J. (2009). *The Quest for Artificial Intelligence*. Cambridge University Press.

Nir, Y., & Tononi, G. (2010). Dreaming and the brain: from phenomenology to neurophysiology. *Trends in Cognitive Sciences*, 14(2), 88–100.

Ohayon, M. M., Guilleminault, C., Zulley, J., Palombini, L., & Raab, H. (1999). Validation of the sleep-EVAL system against clinical assessments of sleep disorders and polysomnographic data. *Sleep*, 22(7), 925–30.

Orlin, J. (2013). Sleep Tracking Startup Zeo Says Goodnight. Retrieved from http://techcrunch.com/2013/05/22/sleep-tracking-startup-zeo-says-goodnight/

Oropesa, E., Cycon, H. L., & Jobert, M. (1999). Sleep Stage Classification using Wavelet Transform and Neural Network. *International Computer Science Institute.*

Park, H., Pa, K., & Jmn, D. (2000). Hybrid neural-network and rule-based expert system for automatic sleep stage scoring. In *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 1316–1319.

Penzel, T., Hirshkowitz, M., Harsh, J., Chervin, R. D., Butkov, N., Kryger, M., … Chesson, A. L. (2007). Digital analysis and technical specifications. *Journal of Clinical Sleep Medicine,* 3, 109–120.

Popovic, D., Khoo, M., & Westbrook, P. (2014). Automatic scoring of sleep stages and cortical arousals using two electrodes on the forehead: Validation in healthy adults. *Journal of Sleep Research*, 23(2), 211–221.

Rama, A. N., & Zachariah, R. (2013). Normal Human Sleep. In *Encyclopedia of Sleep*, 16–23).

Rechtschaffen, A., & Kales, A. (1968). A manual of standardized techniques and scoring system for sleep stages of human subjects. *Washington, D.C.: U.S. Government Printing Office*, *NIH Public*.

Robert, C., Guilpin, C., & Limoge, A. (1998). Review of neural network applications in sleep research. *Journal of Neuroscience Methods*, 79(2), 187–193.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536.

Russell, S. and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. *Prentice Hall, Englewood Cliffs, NJ*.

Sack, R. L., Auckley, D., Auger, R. R., Carskadon, M. A., Wright, K. P., Vitiello, M. V, & Zhdanova, I. V. (2007). Circadian Rhythm Sleep Disorders: Part I, Basic Principles, Shift Work and Jet Lag Disorders An American Academy of Sleep Medicine Review. *Sleep*, *30*(11), 1460–1483.

Sadeh, A., & Acebo, C. (2002). The role of actigraphy in sleep medicine. *Sleep Medicine Reviews*, 6(2), 113–24.

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229.

Schimicek, P., Zeitlhofer, J., Anderer, P., & Saletu, B. (1994). Automatic sleep-spindle detection procedure: aspects of reliability and validity. *Clinical EEG (Electroencephalography)*, 25, 26–29.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117.

Schredl, M. (2013). Nightmare frequency in a representative German sample. *International Journal of Dream Research*, 6(2), 49–52.

Schulz, H. (2008). Rethinking sleep analysis. *Journal of Clinical Sleep Medicine,* 4(2), 99–103.

Schürmann, T., Weeß, H., Binder, R., Steinberg, R., Scchredl, M., Hofmann, M., … Todorova-Rudolph, A. (2001). Validierung des Landecker Inventar zur Erfassung von Schlafstörungen. *Somnologie*, 5(2), 50.

Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. M. (Eds.). (2012). *STOP, THAT and One Hundred Other Sleep Scales*. New York, NY: Springer New York.

Shambroom, J. R., Fábregas, S. E., & Johnstone, J. (2012). Validation of an automated wireless system to monitor sleep in healthy adults. *Journal of Sleep Research*, 21(2), 221–230.

Sinha, R. K. (2008). Artificial neural network and wavelet based automated detection of sleep spindles, REM sleep and wake states. *Journal of Medical Systems*, 32(4), 291–299.

Slater, J. A., Botsis, T., Walsh, J., King, S., Straker, L. M., & Eastwood, P. R. (2015). Assessing sleep using hip and wrist actigraphy. *Sleep and Biological Rhythms*, 13(2), 172–180.

Stanford University (2016). Artificial Intelligence and Life in 2030. *ACM SIGPLAN Notices*, *27*(5), 52.

Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, 59, 433–460.

Van De Water, A. T. M., Holmes, A., & Hurley, D. A. (2011). Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography - a systematic review. *Journal of Sleep Research*, 20(2), 183–200.

Ventouras, E. M., Economou, N., Kritikou, I., Tsekou, H., Paparrigopoulos, T. J., & Ktonas, P. Y. (2012). Performance evaluation of an Artificial Neural Network automatic spindle detection system. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 4328–4331.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, 427(6972), 352–355.

Walker, R., & Russo, V. (2004). Memory consolidation and forgetting during sleep: A neural network model. *Neural Processing Letters*, 19(2), 147–156.

Wallant, D. C. t., Maquet, P., & Phillips, C. (2016). Sleep Spindles as an Electrographic Element: Description and Automatic Detection Methods. *Neural Plasticity*.

Waterman, D., Elton, M., & Kenemans, J. L. (1993). Methodological issues affecting the collection of dreams. *Journal of Sleep Research*, 2(1), 8–12.

Williams, D. L., MacLean, A. W., & Cairns, J. (1983). Dose-response effects of ethanol on the sleep of young women. *Journal of Studies on Alcohol*, 44(3), 515–23.

Yegnanarayana, B. (2009). *Artificial neural networks*. PHI Learning Pvt. Ltd.

Zinkhan, M., Berger, K., Hense, S., Nagel, M., Obst, A., Koch, B., … Stang, A. (2014). Agreement of different methods for assessing sleep characteristics: a comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography. *Sleep Medicine*, 15(9), 1107–1114.

Appendices

Table of appendices

(wird noch ergänzt)

Erklärung über die Eigenständigkeit der erbrachten wissenschaftlichen Leistung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich / unentgeltlich geholfen.

1. Herr Leugering: Maßgebliche elektrotechnische Entwicklung der PCB, anfangs gemeinsam mit Prof. Dr. Pipa, Programmierung des Bluetoothchips und des Microcontrollers sowie eines rudimentären Demo-Skriptes, um die Bluetoothverbindung herzustellen, technischer Teil der Bestellung der PCB.

2. Studierende, die ihre Bachelorarbeit im Themenkomplex Traumschreiber-System geschrieben haben und deren Ergebnisse in dieser Dissertation verwendet wurden (Bachelorarbeiten im Online-Appendix). Alle wurden von mir betreut, die Aufgabenteilung war wie folgt:

- Erste Evaluierungsstudie
  - Kristoffer Appel: Studienidee, Ethikantrag inkl. Studiendesign, Herstellung der Schlafmaske und Programmierung des Experimentes, Wiederholung der Datenanalyse, Einarbeitung und Betreuung des Studenten,
  - Frederik Nienhaus: Versuchspersonenrekrutierung, operative Durchführung der Studie (d. h. Datensammlung im Schlaflabor), Datenanalyse (Ergebnisse nicht verwendet in dieser Dissertation).
- Zweite Evaluierungsstudie
  - Kristoffer Appel: Studienidee, Ethikantrag inkl. Studiendesign, Herstellung der Schlafmaske und Programmierung des Experimentes, Einarbeitung und Betreuung der Studentin,
  - Laura Mandt: Versuchspersonenrekrutierung, Erstellung des “Fragebogen zum Gebrauch des Traumschreibers” in Rücksprache mit mir, Aufnahme des ersten Instruktionsvideos, operative Durchführung der Studie (d. h. Aushändigen und Einsammeln des Traumschreiber-Systems, statistische Datenanalyse in Rücksprache mit mir).
- Explorative Android-App-Programmierung
  - Kristoffer Appel: Betreuung des Studenten,
  - Martin Jäkel: Programmierung der App.
- Literaturrecherche zur Stimulation des Schlafes
  - Kristoffer Appel: Studienidee, Betreuung der Studierenden,
  - Sophia Braumann: Literaturrecherche, Einbringen eigener Experimentideen.

Weitere Personen waren an der inhaltlichen materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

(Ort, Datum) (Unterschrift)

Technical specifications of the artificial neural networks used in this dissertation

Please note, that it was not the aim of this dissertation to optimize the performance of the sleep staging algorithm, but only to form a basis, which future implementations can use as a starting point. This means, that the technological foundation (Keras/tensorflow deep neural networks running on the Raspberry Pi minicomputer) is there and functions in real-time.

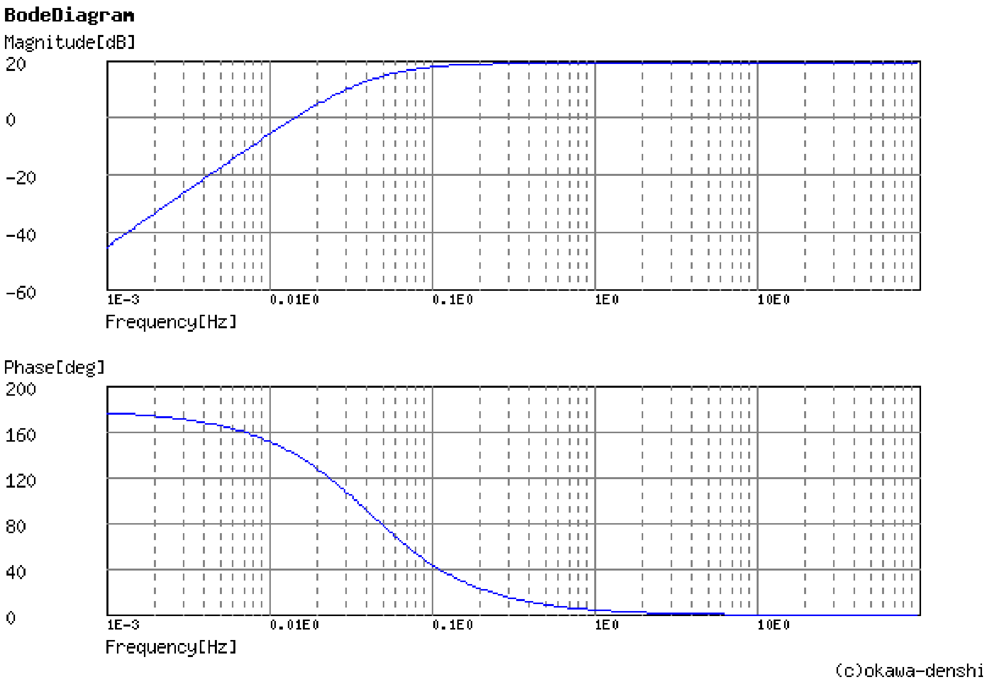
Technical specifications of the artificial neural network for sleep staging

A neural network has been trained to stage sleep data of seven subjects (one night each), recorded with the Traumschreiber sleep mask. The network expects twelve floats as inputs – which are based on the filtered signal of the three recorded EEG channels of the suggested channel configuration, which is then transformed via a Fast Fourier Transform, and finally four frequency bands (0-4 Hz, 5-8 Hz, 9-16 Hz, 25-35 Hz) of each of the three EEG channels are used as neural network input (i. e. descriptors of the second to be classified). The neural network consists of a sequential model of only four layers: after the input layer of 12 neurons, there is a layer of 64 rectified linear units fully connected to the input layer, followed by a dropout layer with a dropout rate of 0.5, and again a fully connected layer of 5 neurons with a softmax activation function, acting as the output layer with one neuron for each sleep stage. The model was trained using categorical crossentropy as loss function with the adam optimizer and accuracy as metric, using n-fold cross validation with a validation split of 0.33, a batch size of 128, during 800 training epochs, and tested on a separate subset of the data. Training took several hours on a single computer with an octa-core i7 processor (not trained on graphics card). The training and testing performance (accuracy) was 0.579 and 0.580, respectively, meaning that the model classified 58% of the (unseen) seconds during testing correctly. This performance is much worse than what modern classifiers can achieve (using the same underlying technology, about 90% accuracy). However, note that the classification here is done based on single seconds (not 30 sec epochs) and on only four frequency bands, i. e. using a very simplistic approach. Again, it was not the goal of this dissertation to optimize the classification performance, but only to set the technological basis, enabling more sophisticated further investigations with possibly more training data in the future.

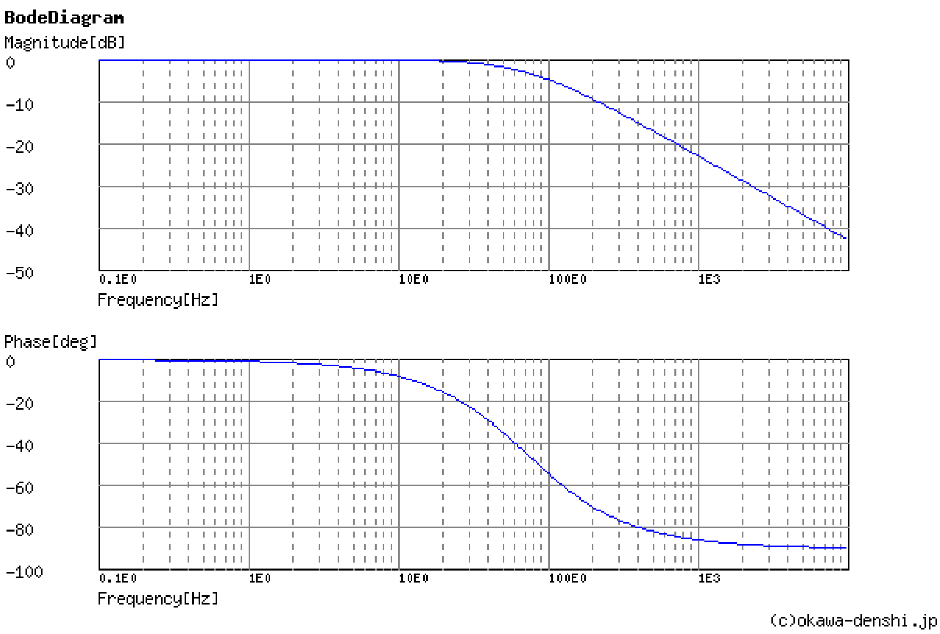
Technical specifications of the using artificial neural network for real-time pattern detection

A neural network has been trained to identify specific eye movements in a real-time recording of the *Traumschreiber* sleep mask. The network's task was to differentiate, whether a left-right-left-right eye movement (LRLR), LRLRLR (3 times), LRLRLRLR (4 times) or none of these was conducted by the subject (the author of this dissertation) during a five second interval. Training data were generated by using an experiment xml file, which instructed the minicomputer to record data and simultaneously present in total 99 auditory stimuli ('2', '3', '4'), upon which the subject was asked to move the eyes accordingly twice, three times or four times to the left and right. The time stamps of the stimuli were then used for extracting five second bins of training data *after* each stimulus onset for each of the three types of stimuli. The five second bins of data *before* each stimulus onset were used for the 'none' category. The network expects 1220 floats as inputs (5 seconds @ 244 Hz sampling rate) – which are based on the median and bandpass filtered signal of the HEOG channel of the suggested channel configuration. The neural network consists of a sequential model of ten layers: after the input layer of 1220 neurons, there are several rectified linear unit convolution and max pooling layers, followed by a dropout layer with a dropout rate of 0.5, and a fully connected layer of 5 neurons with a softmax activation function, acting as the output layer with one neuron for each pattern type ('none', '1', '2', '3', '4'; '1' not being trained). The model was trained using categorical crossentropy as loss function with the adam optimizer and accuracy as metric, using n-fold cross validation with a validation split of 0.33, a batch size of 256, during 4000 training epochs, and tested on a separate subset of the data. Training took about 20 minutes on a single computer with an octa-core i7 processor (not trained on graphics card). The training and testing performance (accuracy) was 1.0 and 0.993, respectively, meaning that the model classified 99% of the (unseen) patterns during testing correctly. Note, however, that only data of one person (the author of this dissertation) was used for training and that, thus, the network might not be able to generalize welll. Again, it was not the goal of this dissertation to optimize the classification performance, but only to set the technological basis, enabling more sophisticated further investigations with possibly more training data in the future.

Filter responses of the analog filters in the Traumschreiber sleep mask



High pass filter response



Low pass filter response