Mobile Health Technology for Personalized Behavioral Activation Therapy

Rohani, Darius Adam

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Ph.D. Thesis
Doctor of Philosophy

cachet | Copenhagen Center for Health Technology

DTU Health Tech
Department of Health Technology

Mobile Health Technology for Personalized Behavioral Activation Therapy

Darius Adam Rohani

Kongens Lyngby 2019

DTU

DTU Health Tech
Department of Health Technology
Technical University of Denmark

Ørsteds Plads
Building 345C
2800 Kongens Lyngby, Denmark
Phone +45 4525 3031
healthtech-info@dtu.dk
www.healthtech.dtu.dk

Summary

In-active behavior, lack of interest in activities, constant feeling of hopelessness, and guilt characterize depressive symptoms. Twenty percent experience depressive symptoms during their life and only 1 out of 4 receive proper treatment.

Behavioral Activation (BA) is a simple therapy method for depressive symptoms. It focuses solely on behavior, in order to support planning and enacting in pleasant activities – moving your body moves your mind.

Smartphones are ubiquitous sensing devices capable of capturing the behavior of their owner. Therefore, they can be a prime tool for BA therapy. We investigated how smartphones can be used in BA therapy. First, we showed that sensing of vigorous activity, time at home, and screen-on duration are reliable features to detect depressive symptoms. Then, we teamed up with the Psychiatric Center Copenhagen, Rigshospitalet and patients to develop a smartphone app called `Moribus` to assist the clinic's current BA therapy program. Instead of registering activities and mood symptoms on paper, they were able to register on `Moribus` and receive visual analytic feedback. We ran a clinical feasibility study on eight of their BA patients. Overall, they stated that they preferred `Moribus` and gained valuable insights from the visual analytic tool.

`Moribus` was able to accelerate therapy sessions, but only for the "lucky few" patients that received treatment. To provide a tool independent of the clinic, we used data from `Moribus` in order to build a recommender model that can suggest specific activities personalized from past behavior. We ran a new design process with patients to incorporate the model in an app: `MUBS`. We ran an 8-week long study of `MUBS` on 17 patients. The patients that used `MUBS` had a significant reduction in depressive symptoms. The app helped them become aware of the smaller positive things in life and enabled them to plan more pleasant activities.

The PhD thesis showed that computational tools can be a perfect medium to assist patients with depressive symptoms to achieve better well-being.

Resumé (Danish)

Depressive symptomer er karakteriseret ved inaktiv adfærd, mangel på interesse, konstant følelse af håbløshed og skyldfølelse. Gennem livet oplever 20 procent af alle mennesker depressive symptomer, og kun 1 ud af 4 modtager korrekt behandling.

Adfærdsaktivering (AA) er en terapiform, som udelukkende fokuserer på adfærd. Det overordnede formål er at få den enkelte person til at planlægge og udføre lystbetonede aktiviteter – bevægelse af kroppen bevæger psyken.

Smartphones er spækket med sensorer, som gør dem i stand til at måle en brugers adfærd. Dette gør smartphones til et oplagt værktøj til brug for AA terapi.

I denne PhD afhandling har vi undersøgt, hvordan man kan bruge smartphones til AA terapi. Først viste vi, at der var en pålidelig sammenhæng mellem depressive symptomer og hhv. hård fysisk aktivitet, tiden du tilbringer hjemme og mængden af smartphone brug. Derefter tog vi kontakt til Psykiatrisk Center København på Rigshospitalet, og gennem et såkaldt brugercentreret design har vi sammen med patienter og klinikere skabt appen `Moribus`. Appen har til formål at assistere klinikkens nuværende AA terapiforløb. I stedet for at registrere aktiviteter og stemningsleje på papir, kan det nu gøres i appen. `Moribus` blev testet i et klinisk pilotstudie med otte patienter. Overordnet set foretrak de `Moribus`, og de visuelle analyseværktøjer i appen gav dem værdifuld indsigt i deres adfærd.

`Moribus` gør det muligt at forbedre terapi, men det er stadig kun de heldige, som allerede er indskrevet, der får hjælp. Vi brugte aktivitetsdata fra `Moribus` til at bygge en anbefalingsmodel, der kan komme med forslag til specifikke aktiviteter personliggjort ud fra tidligere registrerede aktiviteter. En automatisk anbefalingsmodel vil gør det muligt at tilbyde AA uden om klinikken. Vi kørte et nyt designforløb med patienter for at integrere anbefalingsmodellen i en ny app: `MUBS`. Vi testede `MUBS` i et 8 ugers langt klinisk forsøg på 17 patienter. Patienterne der brugte `MUBS` havde en signifikant reduktion i depressive symptomer. `MUBS` gjorde, dem opmærksomme på de helt små positive aktiviteter i livet. Derudover fik `MUBS` dem til at planlægge og udføre flere lystbetonede aktiviteter i hverdagen.

PhD afhandlingen viser, at digitale værktøjer er vejen frem for at hjælpe patienter med depressive symptomer til at opnå en bedre trivsel.

Preface

This Ph.D. thesis was prepared at the Department of Health Technology at the Technical University of Denmark in fulfillment of the requirements for acquiring a Ph.D. degree in Assistive Technologies for Mental Health. The thesis was done under supervision of Professor Jakob E. Bardram, and Clinical Professor MD DMsc Lars V. Kessing, and includes 5 published scientific articles, 2 scientific articles in submission, and 2 popular scientific articles.

The thesis begins a chapter with an historical perspective. This serves the purpose of directing the audience attention towards the importance of the chapters theme. I hope that the audience will find this dissertation as enlighten as the PhD project has been for me. To be responsible within all aspects from developing, to evaluating a tool for patients experiencing depressive symptoms, has been a life transcending journey. This covers everything from organizing workshops with patients, designing a prototype, program it, develop and embed theoretical models, enroll patients, conduct end-interviews, and use qualitative and quantitative analysis methods to understand use and clinical outcome. Therefore, I would love to indulge in discussion within all these aspects with any interested reader.

Kongens Lyngby, November 15, 2019

Darius Adam Rohani

Acknowledgements

I would like to thank my supervisor, mentor, and inspiration, Professor Jakob Eyvind Bardram. As a professor that knows how to chug beer, I was amazed how visionary brilliant you were at grasping the bigger picture. I am thankful for our hour long brainstorms, design meetings, and the support in writing academic literature.

I would like to thank psychologist Nanna Tuxen, and the entire clinical design team, co-supervisor Lars Vedel Kessing, Maria Faurholt-Jepsen, and Mads Frost for the support and discussions we had in early work on `Moribus` & `MUBS`. An extra thanks to Nanna Tuxen for the countless meetings I spontaneously requested to perfectionize design considerations, and for the PhD video interview.

Thanks to nurses Bente Nørgaard Støyer, and Ida Palmblad Sarauw-Nielsen for finding interested patients for our studies. Clinical secretary Helen Gerdrup Nielsen, for helping me find and unlock unoccupied rooms to conduct interviews with the patients.

A thanks to Associate Professor Per Bækgaard for his design inputs, and allowing me to 'hijack' his entire UX-course students for a 4-week trial of `Moribus`.

A special thanks to Andrea Quemada Lopategui for being an immensely big support during my PhD. For the tremendous work on the design that the patients can now benefit from. Your help to arrange the workshops, and the end interview questions. For all our fun moments, laughter's, frustrations (mostly due to Xamarin code, lack of milk for the coffee, and reviewer 2). Your energy and smile has been a source of light stronger than any type of therapy.

A thanks to Professor Steve Whittaker, Aaron Springer, and the research team at University of California, Santa Cruz USA for hosting me on my external stay. It was truly an inspiration to be part of the research meetings, and our well-written paper was an excellent example of that. Steve Whittaker and Marilyn Walker, I am so happy that you introduced Andrea, and me, for the Tom Kha Gai dish.

A big thanks to Adnan Vilic for showing me what 'Applied' truly means. I will never forgot how you helped me debug Xamarin code to make the SQLite database work for `Moribus`. For always being helpful throughout the PhD, and for your time to look through this dissertation with a second pair of eyes.

I would like to thank the entire CACHET team, fellow PhD students, student assistance, and programmers for lifting the spirit, and making these three years a wonderful and memorable chapter in my life. Thanks to PhD students Raju and Devender for our desk-to-desk laughs and discussions. Devender, I did not believe in

20h work-days until I met you. Giovanna for introducing me to the qualitative side of research, and postdocs Steven and Alban, for showing me that postdoc life can be fun and joyful. Student workers, Maja, Thomas, Barnabas, Lori, Julia, Mads, and Björgvin. It was inspiring to see your joy working with health care technology.

I am grateful for Innovation Fund Denmark, and CACHET for funding this research, and Augustinus Fonden for supporting my external stay in Santa Cruz, USA.

I am really thankful for the patients and participants in the study. A large part of this research is contributed and shared to fellow researchers as a product of your dedication and help.

Last, a sincerely thanks to my family, mom, dad, and friends - outside of academia - for supporting me with great dinners, necessary leisure time, but most importantly for the understanding and lack of abandonment as I did not dedicate as much time to you, as I did to my PhD.

Contents

Acronyms

AC Activity Category. xv, xvi, 6–9, 24, 25, 30, 31, 33, 42, 46, 53–56, 71–74

ATUS American Time Use Survey. 71

AWS Amazon Web Services. 46, 57

BA Behavioral Activation. i, xv, 2–11, 14–21, 23, 24, 30, 32, 35, 37, 38, 41, 44–46, 56, 59, 61–64, 66, 69, 70, 72, 73, 75–77

BDI Beck Depression Inventory. 15, 20

CBT Cognitive Behavioral Therapy. xv, 2, 3, 13–16, 18, 19, 24

CUMACF CACHET Unified Method for Assessment of Clinical Feasibility. xvii, 61, 62

DSM-5 Diagnostic and Statistical Manual, fifth edition. 14

FC Feature Category. xix, 26, 28

GCP Google Cloud Platform. xix, 52, 53, 55, 57

HCI Human-Computer Interaction. 3, 6, 7, 36, 37, 39, 44, 45, 59

HRSD Hamilton Rating Scale for Depression. 15

IFTTT If This Then That. 43

LME Linear Mixed-Effects. xix, 63, 65

MC Multiple Comparisons. 25, 30

mHealth mobile Health-related. 2–9, 11, 13, 17, 18, 20, 21, 24, 32, 35–38, 56, 59, 69–71, 73, 76, 77

ML Machine Learning. 29, 31

MVVM Model-View-ViewModel. 49

NB Naive Bayes. xvi, 6, 8–10, 30–33, 44, 51, 53, 69

PCD Patient-Clinician-Designer. 39

PES Pleasant Event Schedule. 72

PHQ Patient Health Questionnaire. xvii, xix, 18, 21, 53, 61, 63, 65

PICO Patient problem Intervention, Comparison, and Outcome. 6

PRISMA Preferred Reporting Items for Systematic reviews and Meta-Analyses. 6

PSSUQ Post-Study System Usability Questionnaire. 61

RCT Randomized Controlled Trial. 4, 16, 18–20, 24, 37, 45

S3 Simple Storage Service. 46

SAP Shared Asset Project. 47

SVM Support Vector Machine. 6, 20, 30, 31

UCD User-Centered Design. 4, 6, 11, 39, 43, 44, 56

UI User Interface. 43, 44, 48, 49, 55, 61

UX User Experience. 4, 6, 36, 44, 70

WHO-5 WHO-5 Well-being Index. 53, 61

List of Figures

List of Tables

CHAPTER 1

Introduction

The act of cutting a vein open to let blood flow out, known as bloodletting, has been the longest practiced medical procedure in history. Bloodletting was believed to cure everything from migraine to smallpox and was a first-choice of treatment for centuries [78]. Today, targeted medical procedures for a majority of conditions have replaced bloodletting to offer more personalized treatments to each patient. Unfortunately, in the domain of mental health where there is a wide range of disorders, the first-line of treatment is pharmacotherapy. This practice draws dangerous parallels to the history of bloodletting, as the induced chemical drugs are affecting the brain in all sorts of ways and their effect is not completely understood.

• • •

1.1 Context and Motivation

Depression is a mental health condition characterized by week-long episode of changes in mood, interests and pleasure [149]. The changes involve depressed mood, and a lack of interest and pleasure in things that are normally enjoyable. Depressive episodes can be accompanied with one or more episodes of mania in which case it is defined as a bipolar disorder. Both conditions are known by the umbrella term affective- (or mood-) disorder [43]. Lifetime prevalence of depression and depressive symptoms is 20% and is projected for the next 10 years as one of the leading causes of disability and disease burden worldwide [123, 99]. This includes both (i) *societal*, with lost productivity among US workers estimated to cost the economy 500 billion a year [27], and (ii) *interpersonal* costs with increased risk of severe disorders such as diabetes, cancer, and Alzheimer [153] and accounting for 50% of all suicides worldwide [147].

Treatment options for depression consist of antidepressant medication (or more broadly; pharmacotherapy), psychotherapy, or a combination [149]. Pharmacotherapy, is the first-line of treatment for patients due to its long history of documented efficacy [4], and easy accessibility in primary care [161]. However, the drug effect is first seen after 2 weeks of use [156], while side-effects, usually sexual dysfunction, weight gain, and sleep disturbance, occurs throughout the period of drug use [66]. Furthermore, as found by a large multi-center study [93], the chances that the first type or dosage of prescribed drugs will have an effect on the depressive symptoms is

low (36%). Usually several years with careful monitoring of symptom response and adverse effects is needed before an optimal drug type and dosage is found [70].

Psychotherapy is done with a trained therapist (e.g., psychologist or psychiatrist) in either face-to-face sessions with the patient or group sessions with multiple patients. Cognitive Behavioral Therapy (CBT) is the most used psychotherapy method within a majority of mental disorders, including depression and depressive symptoms. The reason lies in its short-term consultations, problem-solving technique and clinical efficacy [125, 86]. As a drug-free treatment with no adverse effects, CBT has repeatedly been demonstrated to show similar effects to pharmacotherapy [149]. Further, follow-up effects are evident even a year post-treatment [88]. Despite the evidence-based findings, pharmacotherapy is still found as the most used treatment [161]. Face-to-face CBT requires a big pool of qualified and trained clinical staff. The consultations are time-consuming, provided at a high cost, and accessibility is low [131]. Several improvements have been investigated as a result of these barriers. First, Jacobsen et al. [92] showed that the behavioral component of CBT, what he called Behavioral Activation (BA) for depression had the same outcomes as the full CBT. BA is a more straightforward approach, requiring less consultation time, and can be delivered by nurses or junior clinical staff. Second, Linde et al. [110] demonstrated that remote psychotherapy, in the form of video and phone call consultations, or self-help guided instructions (four or less therapy sessions), was comparable to face-to-face psychotherapy. Third, several research groups have used the ubiquitous nature of smartphones with its processing power and embedded sensors to assist in three primary fields within affective disorders; (i) Treatment or intervention, (ii) self- tracking and management, and (iii) Diagnosis [165].

These mobile Health-related (mHealth) systems have been mentioned as the "*the Swiss army knife of the 21st century*" [189, pp. 269] due to eight prominent advantageous: Capacity, 24/7 availability, Equity, immediate support, anonymity, tailored approach, system linking, and lower cost [146, Table 1]. However, after a decade of smartphone-based studies on affective disorders, we are still juggling in an unsettled landscape with minor impact. An unsettled landscape that now counts more than 10,000 mental health apps available in digital distribution platforms [185, 183]. Although patients are expressing the interest in using these apps to support their mental disorder [111], less than 10% do so [188]. Recent initiatives and taxonomies have started to assist patients and clinicians in the navigation of these apps [85]. Additionally, in an act of desperation, researchers have provided re-conceptualizations to guide and support future mHealth system development [132, 152, 165]. They mention, (i) the importance of personalization [172]. A majority of the designed interventions are based on simple self-help documentations easily accessible through the app, or other one-way methods where the patient is self-managing and has to take action without direct guidance. (ii) To validate the mHealth system on patients with a metal disorder [165]. Several large studies have been conducted using participants with no prior history of any mental disorder, i.e., non-clinical samples. This is of concern since major differences in psychosocial functioning exists between non-clinical samples and patients with affective disorder [42]. A difference that is evident

even after two years of symptom free behavior [42]. Furthermore, behavioral patterns as measured through smartphone sensors or other wearable, were also different between the two groups [A.2]. (iii) To utilize Human-Computer Interaction (HCI) research principles [83]. In the last decade only few ($16/139 \sim 11.5\%$) HCI research studies were conducted as clinical feasibility studies [165]. The inherent differences in conducting research between HCI and clinical researcher proves troublesome [29]. However, adopting HCI principles within health has been demonstrated as an important asset in the design of mobile systems [15]. This includes the involvement of clinicians and clinical patients throughout the design and development process. A fruitful interdisciplinary collaboration can achieve better engagement, contribute to refining designs to improve effectiveness and acquire data to better understand user-engagement and experience [100, 165].

The fact that depression is projected as the leading cause of disability, that currently experience a treatment gap of 56.3% [102] has called the attention of a larger Commission constituting The World Psychiatric Association and Lancet Psychiatry [26]. In a recent report, they conclude that within the next decade a major penetration of computer technology into psychiatric clinic is needed to turn the tide. They refer to this as 'Digital Psychiatry'. Therefore, incorporating the above mentioned suggestion will be a critical step towards successful 'Digital Psychiatry'.

In this dissertation we followed the recommended guidelines in the design, development and evaluation of two mHealth app systems. The systems support BA therapy by facilitating activity- monitoring, scheduling, and subjective ratings of depression, and activity outcomes. In the first system, `Moribus`, the patients could access visual relations between their behavior and depressive symptoms to be shared with their therapist during BA consultations. The second system, `MUBS`, featured an intelligent activity recommendation module to assist and support the patient independent of therapy.

1.2 Problem Statement

BA, a smaller therapy component of CBT, was recently found statistically similar to pharmacotherapy and better than the cognitive component in CBT for treatment of depressive symptoms [50]. BA focus solely on behavior, to plan and enact on pleasant activities. A straightforward method, which makes BA a suitable treatment options for patients in all stages of depression [58]. BA is classified as a low-intensity intervention, with minimal therapy contact [22]. Together with empirical support as a stand-alone treatment option, makes BA a viable option as a guided self-help treatment [176].

As smartphones have been argued as an ideal platform for mHealth [189, 77], it might be a promising candidate to facilitate a guided BA self-help treatment. We defined the main research question of the dissertation as follow.

RQ: *What is the feasibility of mobile and pervasive health technology for the treatment and care of depressive symptoms using a behavioral activation approach?*

In our approach to answer the overall research question, we defined a set of sub-questions. In the mHealth literature on depression there was an predominant interest on the smartphone's capabilities to observe and diagnose as the primary system function [165], utilizing the embedded sensors and internet access. A natural first step was therefore, to investigate the relationship between sensor data and depressive symptoms and what role it plays for BA. Secondly, in order to incorporate the core ingredients of BA [91] into a mHealth system, it was important to understand the technical and User Experience (UX) designs of such a system. Lastly, it was essential to investigate whether the system was used as intended by patients and whether the system was beneficial for the target group. As our target group were patients that experience depressive symptoms, a clinical feasibility study was conducted. Clinical, as it involves 'real' clinically diagnosed patients, in a clinical setting where recruitment was done at a mental health department with clinical personnel involved in all stages of the study. Not to be confused with large scale clinical studies that are defined as Randomized Controlled Trial (RCT)'s with a control group to assess causal relationships. To summarize, the concrete sub-questions that guided the research were as follow:

RQ 1: *What is the relationship between mHealth sensor data and depressive symptoms?*

RQ 2: *What is the technical and UX design of mHealth for Behavioral Activation treatment of depressive symptoms?*

RQ 3: *What is the clinical feasibility of mHealth tools for Behavioral Activation treatment of depressive symptoms?*

1.3 Research Methods

Guided by Bardram's '*Fish model*' [11], the methods applied in the research was balanced between both theoretical and empirical work. We adopted the Triangulation model, as presented by MacKay and Fayard' [116], to outline the range of activities within each work field. The outline is shown in Figure 1.1.

The figure provides an overview on how research from the three domains are contributing each other. As an example, the design of `Moribus` was a result of theoretical work in both prior recorded activity data ('Schema') and literature review on sensor data from people with depressive symptoms ('Sensor'), as well as, User-Centered Design (UCD) based on workshops with patients.

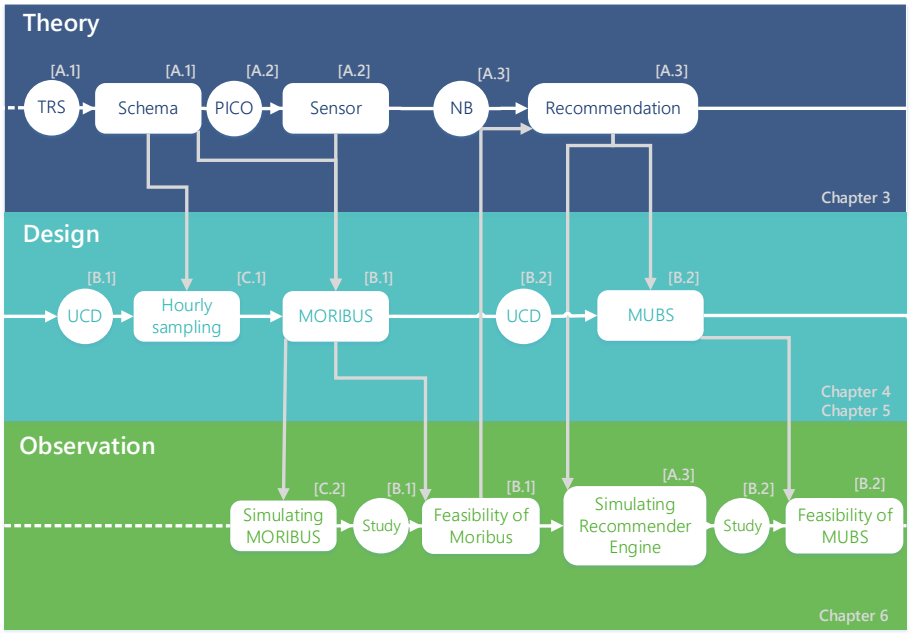


Figure 1.1: An overview of the dissertation. The circular objects represents methods used, while the rectangles are the contributions. The color-coded fields separates the contribution within theory, design and observation, and outlines the dissertation in their respective chapters. * UCD: User-Centered Design, TRS: Transcription of BA activity data, PICO: Patient problem Intervention, Comparison, and Outcome, NB: Naive Bayes, Study: represents the non-randomized feasibility study methodology used.

1.3.1 Theory: Data-driven insights

In line with RQ 1, the overall goal was to understand whether and how collected data from either subjective activity registrations or sensors can facilitate statistical insights for use in mHealth systems. Three distinct activities were carried out as a theoretical foundation for the two mHealth systems. In the first activity, 'Schema', we collected activity schemas, which has been filled in by patients typically over a 4-week period and had been used in BA consultations. The schemas was acquired from the clinical team at the Psychiatric center Copenhagen, Rigshospitalet after receiving informed consent from the patients, and had been filled in following the BA schema filling approach, as instructed by the treatment manual of Lejeuz et al. [104].

Patients were asked to write down their activities on an hourly basis, and score each with a subjective score of the 'Pleasure' (i.e. "*how enjoyable was the activity?*"), and 'Mastery' (i.e., "*the level of perceived accomplishment*"). We transcribed the activities, and categorized it into seven Activity Category (AC)'s. We used a combination of exploratory (post-hoc) data analysis and statistical modelling to infer patterns between activities and the corresponding pleasure score [A.1].

In the second activity, 'Sensor', we followed the Patient problem Intervention, Comparison, and Outcome (PICO) worksheet guidelines [167] and initiated a systematic review of studies that measured the correlation between objective sensor data and subjective reported depressive symptoms. The search was done within ten larger scientific literature databases and limited to studies published between 2007 and 2017, i.e., from the beginning of the smartphone era. We reported our findings using the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) reporting methodology [130].

In the last activity, 'Recommendation', we collected 5,606 registered activities from two separate mHealth user studies [B.1][177]. Hence, the connective line from the empirical work of `Moribus` as seen in Figure 1.1. Contrary to the first activity, data-driven insights were now used within a built probabilistic Naive Bayes (NB) machine-learning algorithm. The NB algorithm was modeled as a recommender system to classify activities as either enjoyable or neutral. In this way the data-driven insights was transformed into actionable activity suggestions. We compared between a generalized model (i.e., pooled all participants together as one entity), and a personalized model (i.e., each user modelled and validated separately), and benchmarked the NB algorithm with a Support Vector Machine (SVM) algorithm, random guessing, and a '*dumb*' classifier that always classifies the activity as neutral.

1.3.2 Design: UX and Technical

In line with RQ 2, the goal was to design and develop mHealth technology that implements a BA approach to assist patients with depressive symptoms. The process was separated into the development of two separate systems. In a first system (`Moribus`), the application was designed to be used adjunct to BA therapy[1]. Particularly, the goal was to replace the handwritten activity schedules, and thereby leverage statistical insights from the digital entries for the convenience of the patient and therapist during consultations. The second system (`MUBS`) was an extension of the first by adding an intelligent module to recommend enjoyable activities personalized to the patient. The intelligent module was designed as a supporting tool that mimics the therapeutic guidance that happens during a consultation.

During development we followed an experimental HCI approach as guided by Hornbæk [89]. For both versions this included (i) an ideation phase in the interdisciplinary team; (ii) an iterative UCD strategy including workshops, and UX design

[1] Where a system is used as a tool for clinicians in their therapy. Also referred to as blended treatment

meetings with both low fidelity paper mockups and high fidelity interactive prototypes; (iii) and an implementation part. For `Moribus` we focused on a design that accommodated exhaustive hourly sampling of activities. We included statistical models from the BA schemas [A.1] in the design, as a visual analytic tool for the patient and therapist. For `MUBS` we integrated the recommender model [A.3] in the design.

1.3.3 Observations: Clinical Studies and Feasibility

In line with RQ 3, the goal was to perform a clinical study in order to evaluate the feasibility of the technology in real patient use, applying both quantitative- and qualitative analysis.

To follow best practice in health-related research within HCI [101] we did a non-randomized single-arm clinical feasibility study of `Moribus` and `MUBS`, recruiting 15 and 28 clinical diagnosed affective disorder patients respectively. The systems were used daily for 4 and 8 weeks, respectively. The mixed method evaluation consisted of a quantitative analysis of usage patterns, and a qualitative analysis from semi-structured interviews and usability questionnaires distributed in the end of the study. The HCI based evaluation helped us understand requirements and context, how patients interact with our applications, and explore the strengths and challenges associated with the different incorporated design features [165]. We ran the study on `MUBS` more than a year after we finished the `Moribus` study. Hence, the findings from the HCI evaluation of `Moribus` were incorporated into the design and evaluation of `MUBS`, as also illustrated in Figure 1.1.

1.4 Research Contribution

The research contributions covers theoretical-, design-, and empirical work. In the theoretical work we developed models to explain how behavioral data relates to depressive symptoms. In the design phase we contribute two mHealth systems utilizing BA techniques. Our empirical work contribute a clinical feasibility study for each system.

1.4.1 Theoretical

During this thesis we developed and published three different theoretical models that describes the relationship between behavior, depressive symptoms, and 'Pleasure'. In the first model we initiated a exploratory data analysis on 2,480 transcribed activities. We used the designed Activity Category (AC)'s [C.1] to label and create statistical models to explain activity data in relation to perceived 'Pleasure'. The model demonstrated personalized patterns within the type of activities done, e.g., P3, contrary to the others, had largest pleasure for social based activities [C.2]. We expanded the model with higher-order statistics [A.1]. For instance, we found that the optimal

amount of social activity was 2 hours for P3. The model on transcribed activities contributed the efficacy of digitalizing BA activity scheduling. It was able to find patterns within the behavioral trait that could assist the therapist and user.

The second model contributed a mathematical model 'weighted directionality of the correlation' to combine the findings of prior work on sensor data and depression [A.2]. The model compares correlation between objective behavioral features collected by mobile and wearable technology, and depressive symptoms as measured by different rating scales. We ran the model on studies that used clinical patients, and studies using non-clinical (i.e., undiagnosed) participants [A.2, Figure 2, 3]. We show that a feature such as the daily duration of the smartphone (i.e., *Screen active duration*) is positively correlated with the degree of depressive symptoms across all studies and on both clinical and non-clinical participants, while a feature such as the amount of send text messages (i.e., *SMS send*) show contradicting correlations for the clinical sample.

The two models contribute to knowledge that can be conveyed to the patient to assist in understanding their behavior and how it affect their mood, what we refer to as *sensemaking*. As a natural extension, we contributed a third model that provides automatic sensemaking. The model was designed as a personalized content-based NB recommender system to suggest enjoyable activities [A.3]. We demonstrated that the model was able to correctly classify an activity as enjoyable or neutral in over 70% of the cases. Furthermore, the personalized version of the model was key to recommend the right activity, as it outperforms a generalized model.

1.4.2 Design

As part of this thesis we designed two mHealth system. In the first system, `Moribus`, we contributed an mHealth app to be used adjunct to BA therapy [B.1]. The system was designed as an hourly-based activity planning and registration tool. `Moribus` featured several design choices to support hourly sampling in BA. We designed six AC labels to yield faster activity scheduling [C.1], and implemented the theoretical model on activity [A.1] as a visual analytic tool in the system. In order for the patient to understand the different AC, we added ten typical example activities for each category. The activities were digitalizing from a list of pleasant activities as generated by the work of Lewinsohn & Macphillamy [118].

In the second system, `MUBS`, we contribute an mHealth app to be used stand-alone as a self-help app that uses principles from BA [B.2]. It follows similar design as `Moribus`, but adds the intelligent content-based recommender model [A.3] to assist the user in planning and performing specific activities. The recommended activities were personalized to the user based on thumbs up and down ratings from completed activities. Novelty was supported by including all activities from Lewinsohn & Macphillamy [118] and expanding with work by Mørch & Rosenberg [135], totalling 384 pleasant activities. To promote planning, we added a 'Difficulty-score'. It represented the effort it takes to initiate the activity. Furthermore, to support a more

balanced lifestyle, we added a progress visualization. The progress was represented as a staircase with three levels for each AC.

1.4.3 Empirically

Empirically we completed two clinical feasibility studies of the systems. The study involved clinical patients diagnosed with affective disorder.

Prior to a clinical evaluation, we ran a simulation study on `Moribus` [C.2]. We used the hand-written BA activity data [A.1] to simulate a normal usage situation. The study demonstrated the feasibility of moving BA activity registration to mHealth technology. We found that the AC's were sufficient to cover all activities. Moreover, the design of the activity registration interface facilitated fast planning and registering of activities.

`Moribus` was evaluated with clinical patients in a 4-week clinical feasibility study, which contribute the feasibility of using the mHealth technology for BA in everyday use by patients [B.1]. The study show that patient found the visual analytic tool useful for learning about their behavior and its effects on mood. Quantitative analysis revealed highly individualized usage patterns. For example, some patients preferred to do the activity registrations in the evenings, while others did it just after an activity was performed.

The activity data from `Moribus` contributed the training of a NB content-based recommender model. The model was subject to a simulation study using data from `Moribus` patients, and the catalog of pleasant activities [A.3]. We simulated the model across the `Moribus` study period. We illustrated how recommended activities from the catalog changed as the patient used the system. The simulation study demonstrated how novelty could be introduced in a recommender algorithm for mHealth technologies like `MUBS`.

In a 8-week long feasibility study of `MUBS` the patient reported high usability and usefulness [B.2]. The patient had a statistical significant reduction in depressive symptoms. Of all the registered activities, over 60% were picked from the catalog of pleasant activities. They found it as an inspirational source to introduce more pleasant activities in their everyday life. Recommended activities were adopted very differently by the patients. It seemed like patients who were full time employed or had a busy schedule in general, did not need recommendations, while patient with more depressive symptoms found the recommendations more useful.

1.5 Scientific Publications in Thesis

In Figure 1.1, each activity is annotated with a citation to relevant scientific publications. This section briefly outline these main publications, which are included as part of the Thesis.

Data-driven learning in high-resolution activity sampling from patients with bipolar depression: Mixed-methods study. In: *JMIR mental health*, 2018 [A.1]

We ran statistical data analyses on 2,480 transcribed activities from patient in BA therapy. The paper contribute the efficacy of digitalizing BA activity scheduling.

Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review. In: *JMIR mHealth and uHealth*, 2018 [A.2]

A systematic literature review was the first of its kind on the correlation between objective behavioral features collected by mobile and wearable technology, and the assessment of depressive symptoms as measured by different rating scales and questionnaires. The central contribution of this work came from combining correlations of the same behavioral features across all studies to present joint visualization for non-clinical and clinical patient.

Recommending Activities for Mental Health and Well-being: Insights from Two User Studies. In: [Submitted] *IEEE Transactions on Emerging Topics in Computing - Special Section on New Frontiers in Computing for Next-Generation Healthcare Systems* , 2019 [A.3]

A probabilistic NB recommender model was developed and tested on 5,605 registered activities from two different populations of non-clinical [177], and clinical [B.1] patients. This research sheds light on recommendations for mental health.

Personalizing Mental Health: A Feasibility Study of a Mobile Behavioral Activation Tool for Depressed Patients. In: *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2019 [B.1]

The paper on `Moribus` explains the entire design, development and evaluation process to assess the feasibility of a activity registration tool to be used adjunct to BA therapy. The evaluation was a mix between qualitative and quantitative results.

Supporting Behavioral Activation in Mental Health by Mobile Activity Recommendations. In: [Submitted] *ACM CHI Conference on Human Factors in Computing Systems*, 2020 [B.2]

The paper on `MUBS` explained the design, development and evaluation process of an BA app to be used independent of the clinic. The evaluation was focused on the qualitative feedback.

1.6 Thesis Overview

The dissertation continues with background information on BA therapy and mobile technology for mental health. The latter features a summary of the state-of-the-art mHealth applications supporting BA. The next four chapters follows the Triangulation model as illustrated in Figure 1.1. Chapter 3 present three different approaches to leverage insights between data and depressive symptoms. Chapter 4 explains how we designed the two mHealth systems. The chapter covers our UCD approach from sketching, and prototyping the systems. Chapter 5 covers the implementation and deployment of the systems. The chapter includes the software architecture of the implemented systems. Chapter 6 describes the clinical feasibility studies that we performed on the developed systems. It explains the study design, recruitment, the process, and findings. Chapter 7 discuss our theoretical, design, and empirical contribution in relation to our presented research questions. The chapter follows with a discussion on the limitations of the dissertation, and suggests future steps. Chapter 8 concludes the dissertation with a summary of the work.

CHAPTER 2

Background and Related Work

In the very existence of life, education - in its various forms - has been a fundamental necessity of life. From the primitive man that was taught to accomplish practicalities such as hunt, use weapons, and secure shelter to the theoretical education of religious ceremony and worship by priests and shamans [134]. Mediation of knowledge has evolved ever since from the first school-group type learning in the Chinese system of education ($\sim$ 1600 B.C.), to Greek education that was the first case of *knowledge for the sake of knowledge* offered to all instead of the privilege of the few [134], to newer history with class-room type of education. Today, the educational system have changed remarkable due to the influence of technology [35]. Not only does it erase geographically and socio-economically issues in accessibility, technology also expanded the possibility to mediate knowledge by different types of media and sources.

• • •

The penetration of technology within psychology and the treatment of depression is in a much earlier phase than the education changes. Now we are starting to see the first examples of technologies being integrated into mental health services; such as the MindSpot clinic in Australia that delivers online CBT treatment for anxiety and depression [181]; and the Monsenso mHealth solution offered at the Centre for Telepsychiatry in the Region of Southern Denmark [37]. Usually they are designed as double-loop systems where patients are doing self-monitoring of their symptoms while the clinical staff have a web-portal to access the data and provide feedback and assistance if deemed necessary [17].

2.1 Behavioral Activation

Regardless of region and gender, the key symptoms of depression are worldwide expressed as: "**Depressed mood/sadness**", "**Problems with sleep**", "**Fatigue, loss of energy**", "social isolation", "**appetite problem**", "crying a lot", "headaches", "**loss of interest**", "general aches and pains" and "**suicidal thoughts**" [82]. The list continues and people who have experienced depressive symptoms usually describe

them as: "*the worst thing they have ever been through, as falling into a deep black hole of mental pain and hopelessness*" [122].

The highlighted symptoms are the ones defined as a diagnostic criterion for depression in the Diagnostic and Statistical Manual, fifth edition (DSM-5). DSM-5 is a widely accepted nomenclature for the classification of depression which states that:

> "*The individual must be experiencing five or more symptoms during the same 2-week period and at least one of the symptoms should be either (1) depressed mood or (2) loss of interest or pleasure*" [9, §MDD.A].

Evidence-based research points at psychotherapy as the efficient choice of treatment [50, 74]. It is based on conversational therapy and the broad definition is given as:

> "*Psychotherapy is the informed and intentional application of clinical methods and interpersonal stances derived from established psychological principles for the purpose of assisting people to modify their behaviors, cognitions, emotions, and/or other personal characteristics in directions that the participants deem desirable*" [199, pp. 218].

It does not involve intake of medication or surgical application, even though alterations in the brain are evident [74]. Psychotherapy covers thousands of different methods, from which CBT is the most used to treat depressive symptoms. It was the the work of Aaron T. Beck supported with empirical evidence for the combination of cognitive therapy and behavioral approaches that changed the mental health services method of treating depression [18]. He formalized a set of clear treatment principles that made the therapy process widely accessible. An example with the principles behind CBT is displayed in Figure 2.1.

It is a problem-solving technique that focuses on the "here-and-now" [18]. Taking the example from Figure 2.1, the therapist would start by identifying the overall goal, and subsequent ask about the current barriers. If the patient is experiencing severe depression they initiate BA principles, as in this example, make smaller activity plans, e.g, create a list of items needed. If the patient is missing some football shoes the therapist would help schedule an activity of buying some new football shoes. In the cognitive therapy the focus surrounds the patients dysfunctional attitudes. In a careful balance of empathizing with the patients emotional experience (e.g., "*I can understand that you feel that the others might be better than you*"), the therapist needs to identify the patients faulty cognition's and steer towards an enhancement of pleasant emotions (e.g., "*What league are they playing? Having an extra player always benefits a team*"). As might be evident from this small example, CBT is a methodology that depends on the skills of the therapists, which again induce cost in terms of training and employment [160]. Jacobson et al. [92] did a component analysis of CBT in order to isolate and compare the BA component against the cognitive component and the full CBT. A sample of 149 clinical participants, were randomly allocated to one of the three treatment groups. Treatment consisted of at least 12

Figure 2.1: The triangular structure displays the principles of CBT. A precise example is given. The red boxes illustrate a common scenario, while the green is a result of re-structuring thoughts behavior.

consultations. They were not able to find any statistical significant difference in depressive outcomes, as measured by Beck Depression Inventory (BDI) and Hamilton Rating Scale for Depression (HRSD), between the three groups. Same observation was seen after a 6-month follow-up. The study was a milestone within psychotherapy since the BA component, is a more straightforward approach that requires less time in the clinic, and can be delivered by junior mental health workers with less intensive and costly training [160].

The origin of BA dates back to the work of Lewinsohn [107]. While Aaron T. Beck with CBT treat cognitive changes (e.g., feeling of guilt, or feeling of social isolation) as the primary causal effect in depression, behavioral theory views these feelings as secondary clarifications of a general unpleasant feeling state (dysphoria); "*I am feeling bad*", or "*I am not likeable*", respectively [107]. Behavioral theory states that depressive behavior of dysphoria, or other somatic symptoms or/and low activation is a result of a low rate of response-contingent positive reinforcements [107].

Positive reinforcements are positive responses or outcomes based on an event or activity that makes the activity more likely to reoccur. Response-contingent indicates that the positive reinforcements is depending on the action of the surrounding environment [122]. For instance the event of "upgrading" your academic status from PhD student to postdoc might result in loss of social reinforcement (e.g., loss of work mates in the transition to a new project). BA thrives on our conception of pleasure and accomplishments. BA treatment aims at increasing exposure to positive consequences of healthy behavior, to increase re-occurrence and thereby reduce the time spent on feeling depressed and carry negative thoughts [105]. The possibility of acquiring positive reinforcements depends on three conditions [107]:

1. The number of potential reinforcements.
2. The available reinforcement from the environment.
3. The acquired educational skill-set to emit behavior that elicit reinforcement.

The most typical treatment procedure, as guided by Martell and Lejuez [105, 122], follows three overall steps for 10-12 sessions each up to an hour long. First, the patient needs to be able to identify reinforcers for depressed and non-depressed behavior. This is done by asking the patient to carefully, during their everyday life, note down - on a paper-based weekly schedule (see [105, Figure 2] for an example) - their activities in detail (hour-by-hour). For each activity they are also told to write a rating (between 1-10) on the perceived 'Pleasure'. This is continued for several weeks to acquire an overview of daily routines. The weekly paper schedules are shared with the therapist and together they locate activities that promote healthy behavior and, in same level of detail, plan next weeks activities accordingly. The instruction for this is simple; follow the plan strictly.

The process of identifying and scheduling activities for the patient is done through a series of tasks. This includes first to understand overall goals and behavior that are valuable to the patient, which may have been present in a previous period of life (i.e., before the depressive symptoms). Activities planned would then be small sub-goals that redirect the patient back to the everyday life that they want to accomplish. Second, rank activities in terms of 'Mastery' or perceived difficulties (see [105, Figure 5] for an example) to understand what activities are easier to initiate and pursue. Last, create a set of sub-goals for each activity including specifying measurable success criteria, such as frequencies. For example, instead of a passive activity such as "*see my brother*", it should be stated in a more active manner like "*see my brother twice a week*".

Figure 2.2 provides an overview of BA therapy and its core ingredients. The ingredients follow the presented treatment outline of Martell and Lejeuz [105, 122], and include those identified by two academic experts and one CBT expert [91].

BA has repeatedly, through large RCT's, found to be as effective as pharmacotherapy and CBT [50], with slower rate of relapse and recurrence of depressive

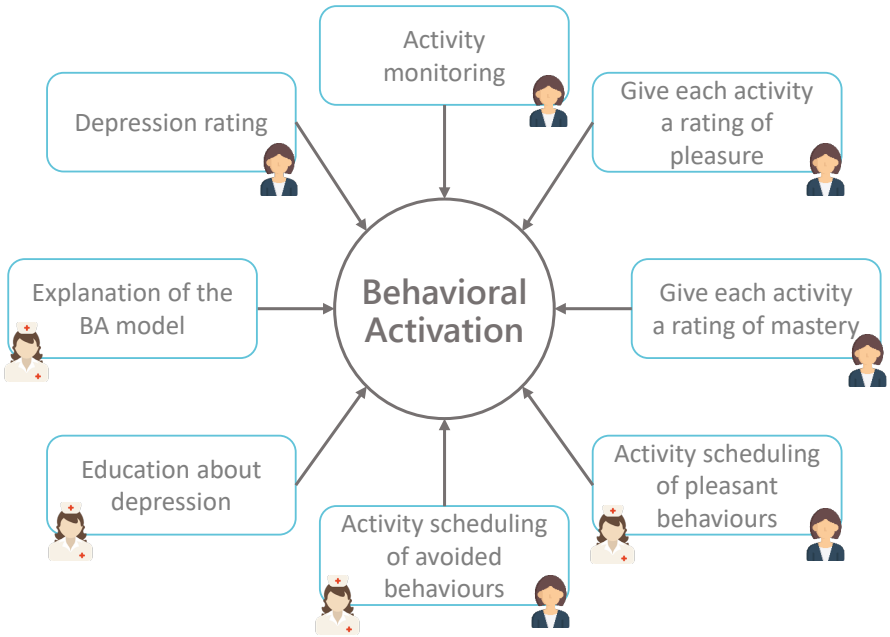


Figure 2.2: The core ingredients in BA therapy. Each action is attached with the appropriate facilitator, either the clinician or the patient.

symptoms [51], and lower therapeutic costs [160]. This, together with the primitive paper-based ways of engaging and motivating patients during their BA treatment (see [105, Figure 8-10] for examples), has set the stage for mHealth technology as a potential good way to explore new approaches to support BA.

2.2 mHealth technology supporting Behavioral Activation

Over 10,000 mHealth applications for mental health exists in the digital distribution platforms [186], and a majority of those are developed in a non-clinical setting without support from research [165]. Several independent organizations have started to index and organize these applications to guide potential users towards high-quality apps suited for their intention. These includes PsyberGuide[1] that is operated by University of California, Irvine and at Feinberg School of Medicine, Northwestern University, and

[1] `https://psyberguide.org/`

MindApps[2] which is a Danish site operated by The Centre for Telepsychiatry in the Region of Southern Denmark.

A fraction of these operate with principles from CBT and BA. A majority (89.73%) of these are not evidence-based, and display low level of adherence to the core ingredients of CBT and BA [91]. The intention of this section is to present and provide the reader with an overview of the most relevant mHealth systems supporting BA, which has been studied and published in a scientific research environment. To do so, we defined three categories, each with two subcategories, to classify the different types of mHealth systems. The categorization is visualized in Figure 2.3, with an example application for each.

As smartphones is within reach 90% of the time [48], the main benefit of mHealth technology is the improved accessibility of support in terms of both providing information and intervention to the patient. Having access to the system 'in-the-moment' can reduce inconsistent recollection of past behavior[3]. Furthermore, the improved ability to scale psychiatric support beyond the normal constrains of time and costs associated with patients meeting their therapist on a regular basis. The last category in Figure 2.3 'Providing BA therapy with mHealth' is defined by its ability to provide actionable feedback to the user. The actionable feedback is technology-enabled and not dependent on clinical involvement. Such systems could be a possible solution to reach patients with stigma towards the clinic. Patients may prefer the anonymity of using a mobile app and feel more comfortable reporting symptoms and behavior, thereby increasing treatment-seeking behavior [184]. In the following sections we present the example apps shown in Figure 2.3, and introduce our own contributions.

2.2.1 Self-educational

Kauer et al. was one of the first to demonstrate the effect of self-monitoring symptoms of depression with a mobile application [97]. In a large RCT with 118 youth and young adults using an app for up to 4 weeks had an statistical significant reduction of depressive symptoms due to, what they mention as, increased emotional self-awareness. Applications within BA emerged much later. `HeadGear` [46] was developed to reduce depressive symptoms through a 30-days challenge plan. The challenges consisted of Behavioral Activation (BA) tasks (i.e., a mindfulness exercise, activity planning, or goal-setting). A single-arm feasibility pilot study with 84 non-clinical participant was conducted for 5-weeks. The participants had a significant reduction in depressive symptoms as measured by the Patient Health Questionnaire (PHQ) 9.

In `Aptívate!` [45] the participants used a daily calendar to schedule and keep track of activities and daily mood. Activities were generated based on identified values within five life areas (e.g., career and education, relationships, and health). Examples of values would be: "*have family dinners twice a week*" or "*call my children*

[2] `https://mindapps.dk`
[3] This is also known as reconstruction bias

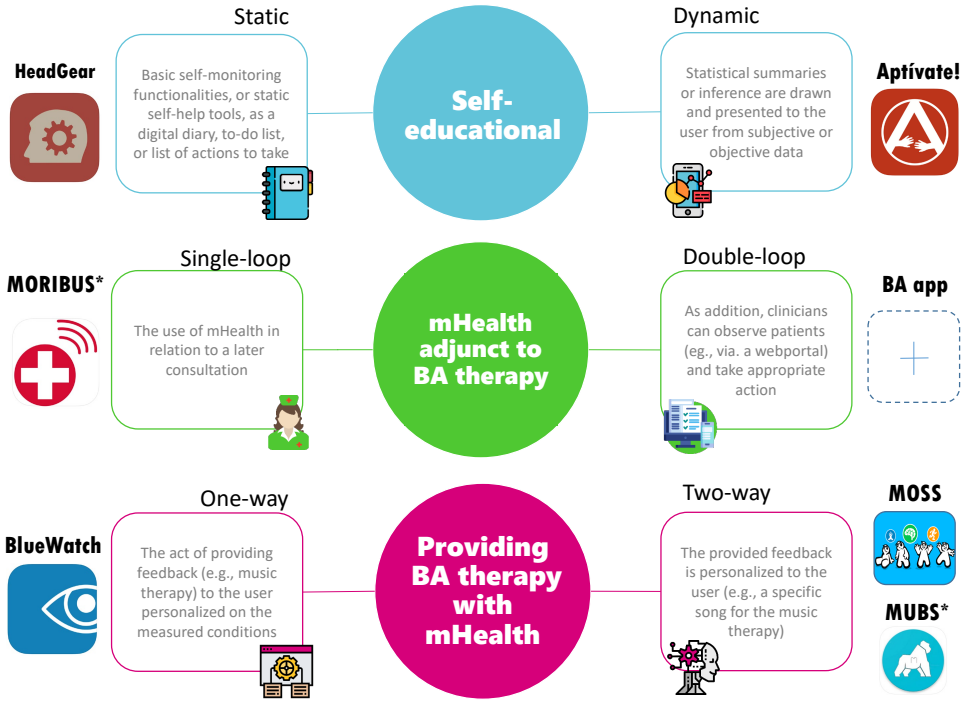


Figure 2.3: mHealth applications categorized into 3 main areas, each with two subcategories. The definition is given in the squared boxes, while example applications are listed in the outer part. *Application contributed in the PhD.

three times a week". A badge reward-system was implemented to motivate continued use. A pilot RCT feasibility study with treatment as usual (i.e., no app) as control was conducted for 8 weeks with 42 non-clinical participants. The participants allocated to `Aptívate!` had significant reduced depressive symptoms over time compared to control, and on average completed 21.73 activities with a 50% retention rate.

2.2.2 mHealth adjunct to BA therapy

Besides the contribution of `Moribus` [B.1], no other single-loop app has been designed to support CBT or BA in therapy. The closest applications, which there is numerous good examples of [190], are related to the Static 'Self-educational' category. Self-monitoring apps have the potential to improve communication between the patient and the clinician during consultations:

> "*Apps help prompt my clients to log things like their mood or whether they exercised or drank alcohol or slept well, so that I can have an accurate picture of their week, as opposed to them trying to reconstruct things on paper the day before therapy or during therapy, which is much less reliable*" [143, pp. 64]

A good example is the `T2 Mood Tracker` app developed by the National Center for Telehealth and Technology, USA [33]. The app was designed to help users monitor and track their health within six predefined areas (anxiety, depression, stress etc.) as well as custom areas such as pain and sleep. These areas contained several opposite adjectives (e.g., tired vs energetic within the depression area) where the user choose the most representative with a continuous slider. The clinicians gather these symptom data between clinic visits to review those data and thereby improve on the consultations.

In the `BA application` the participants could register and reflect on important behaviors in order to increase everyday activation. The app had a database of 54 activities (e.g., "*Read the newspaper*", or "*Take a walk with a friend*"), as well as the support to add new ones [115]. The therapist had access to the data through a web portal, and could send short text messages to the participants. Each week the participant received personal feedback from their therapist. A RCT study compared a full BA treatment consisting of ten face-to-face consultations (n = 46) against the app as adjunct to BA treatment with only four face-to-face sessions (n = 47) [114]. Both condition had a significant improvement in depressive symptoms as measured by BDI. There was no difference between the two conditions, even though the mHealth condition reduced the therapist time with an average of 47%.

2.2.3 Providing BA therapy with mHealth

In 'closed-loop' systems, the user receive actionable feedback directly inside the application. Feedback that would otherwise be given by therapist or equivalent. The feedback is usually completely data-driven and two different methods define the two different subcategories of such systems. Either a One-way or Two-way personalized system. The former usually involves a specific action plan depending on the received data. An example of this is demonstrated in the proposal of the `emHealth` application [198]. They used a decision tree and SVM machine learning algorithm to predict the level of depression (between zero and five) based on questionnaire data within aspects of family life, self-promotion burden, work pressure etc. Dependent on the level of depression a set of feedback possibilities were available. In level 0, the participants receive music therapy. In level 5 the user receive encouragement and praise, with suggestions to cultivate a variety of hobbies [198].

The One-way mHealth system `BlueWatch` was used in a 12-week single-arm usability study (N = 5) [69]. The system was designed as a self-guided app to improve well-being. It used the mood survey data recorded by the patient to unlock specific feedback oriented content. The content included educational material and BA activ-

ities such as creating to-do lists and monitoring daily activities. An example where the provided feedback was personalized to the user was seen in the Two-way `Moss` system [193]. Based on passive sensor data to collect information on time spent at home, the number of steps taken etc. they provided appropriate interventions within four categories (i.e., social activity, physical activity, relaxation, and mindfulness). Furthermore, the activities in each category were personally fitted each user through past ratings and how many times the activity have been cancelled. The categories were presented on the main screen of the app as circles. When a user selected a circle they received the top three personally fitted activities. With an 8-week single-arm feasibility study (N = 12) they showed a statistical significant drop in depressive symptoms as measured by the PHQ scale.

In summary, we can draw following conclusions on the related systems:

- None of the systems satisfy all the eight BA components presented by Huguet et al. [91]. For instance, Wahle et al. [193]'s Two-way system with personalized feedback does not support activity scheduling.
- There is a sparse selection of mHealth systems designed to be used with therapist during consultations. Of the few that exist, the `Ba application` does a good job in activity scheduling and qualitative reflections on past behavior. Unfortunately, the app does not focus on data-driven insights, and provides a single quantitative visualization of the frequency of activities [115, Figure 2].
- Wahle et al. [193] was the only Two-way system to provide personalized feedback to the users. However, the recommended actions - the specific activities - were chosen solely on the user's past rating. The system did not leverage information about the activity, and hence new 'unseen' activities (that carry no information) were never true recommendation. This approach does not follow standard recommendation methodologies such as collaborative- or content-based filtering [59]. In these cases, as an example, an activity would be recommended because either a person that matches your preferences also liked that particular activity, or from past activities the system knows that you enjoy that type of activity in the given context.

To address the gap, the target of the thesis was to design, implement, and evaluate two separate mHealth systems. The first, `Moribus`, to fill the gap within technology adjunct to BA therapy [B.1]. The system features a visual analytic tool with detailed personalized insights into relationships between behavioral patterns and depressive symptoms. It supports both the patient and the therapist as the insights can be leveraged by less-training clinical personnel including nurses and psychology trainees, to assist in BA therapy. The second system, `MUBS`, is able to include those experiencing stigma towards the healthcare system. It is a Two-way system with a content-based recommender system [A.3] that provides personalized feedback to guide the patient towards better well-being [B.2].

CHAPTER 3

Data-driven insights

In the area of Westminister, London in 1854 there was a severe cholera outbreak killing more than 600 people within 10 days [174]. The later investigation by John Snow[1] became a milestone within epidemiology. By collecting data on mortality reports, visited homes of cholera victims, and locations of public water pumps, he was able to create a "topography of the outbreak" [174]. The map substantiated his theory that cholera was water rather than airborne.

> *"Each water company supplied alike both rich and poor, and thus there was a population of 300,000 person of various conditions and occupations, in intimately mixed together, and divided into two groups by no other circumstances than the difference of water supply"* [174, pp. 241]

• • •

Connecting various data types and communicating the insights in a clear way – as done by John Snow's map – is a good example of the possibilities of data-driven insights. Possibilities that are beyond health outcome research, but targets all domains within healthcare [76]. For patient-facing applications data-driven insights are praised for (i) increased patient engagement, (ii) improved treatment adherence rates, (iii) supporting physician-patient communication, and (iv) more effective and efficient service by clinicians [76]. Particularly, within psychology, smartphones has been mentioned as a key technology to facilitate data-driven insights [187] as they poses the possibility of ubiquitously to gather and present behavioral data and self-reported symptoms.

In the following sections, we present our theoretical contributions within data-driven insights to support BA. The theoretical work was either drawn from smartphone-based data, from the literature or data collected ourselves, or data collected from past BA therapy sessions.

[1] An English Physician born in 1813 - not to be confused with later fictional characters

3.1 Statistical methods for real-world BA activity registrations

A large meta-analysis by Ekers et al. included 26 larger RCT studies on the effectiveness of BA [58]. They highlight that BA is comparable to CBT and pharmacotherapy. Furthermore, Torous et al. have argued for the advent of mHealth in psychiatry [187]. However, in a review of 12 BA apps, they found that there was low level of adherence to the core methods of the BA models and questions the utility of these apps [91]. A common ground would be an app, to be used adjunct to therapy, with BA ingredients that naturally adhere to the core guidelines.

Towards the design of such an app, a first step was to develop a theoretical model that translates activity data into statistical insights on behavior. An incentive to move the activity registration part of BA therapy to a digital solution. Not only does a digital solution enable less experienced therapist to leverage these insights towards better treatment, it is also (i) empowering patients in their own behavioral understanding, (ii) less subjective to retrofitting of self-assessment data [15], and (iii) carries functions such as notifications, and auto-filling that can improve adherence [55].

The work has been presented in [A.1]. Briefly, we collected paper-based activity registrations from patients experiencing depressive symptoms. For each activity, the patient registered a 'Pleasure' score. Two independent researchers transcribed the activities to a digital form and labeled each with an AC. We ran several statistical analyses to understand the details of current BA activity sampling and the kind of insight that can be gained.

We collected 2,480 hours of self-reported activity data from seven patients that had received BA therapy at the Psychiatric Center Copenhagen, Rigshospitalet. Among several statistical significant insights derived from the transcribed data, we underline a particular part of our theoretical model as shown in Figure 3.1 For each patient, we took the global median from the daily average 'Pleasure'. The days that had an daily average 'Pleasure' above or below the global median was classified as 'good days', and 'worse days' respectively. For each day (d), we weighted the total hours of each AC (C) with the fraction between the global median ($\tilde{G}$) and the daily average 'Pleasure' (P). The calculation (wDist) is formalized in Equation 3.1.

$$
\begin{aligned}
\text{wDist}_{d,C} &= \sum_{i \in t} A_{d,t,C}(x) \left(\frac{\sum_{i \in t} P_i / |t|}{\tilde{G}} \right) \\
A_{d,i,C}(x) &= \begin{cases} 1 & \text{if } x = C, \\ 0 & \text{otherwise.} \end{cases} \\
t &= [8, 23] \\
C &\subseteq \{\text{movement, work \& edu.,...}\}
\end{aligned} \tag{3.1}
$$

t is the time of day (24h format), while A is the list of activities. All the weighted distances (wDist) of each AC for each patient was then subject to a nonparametric

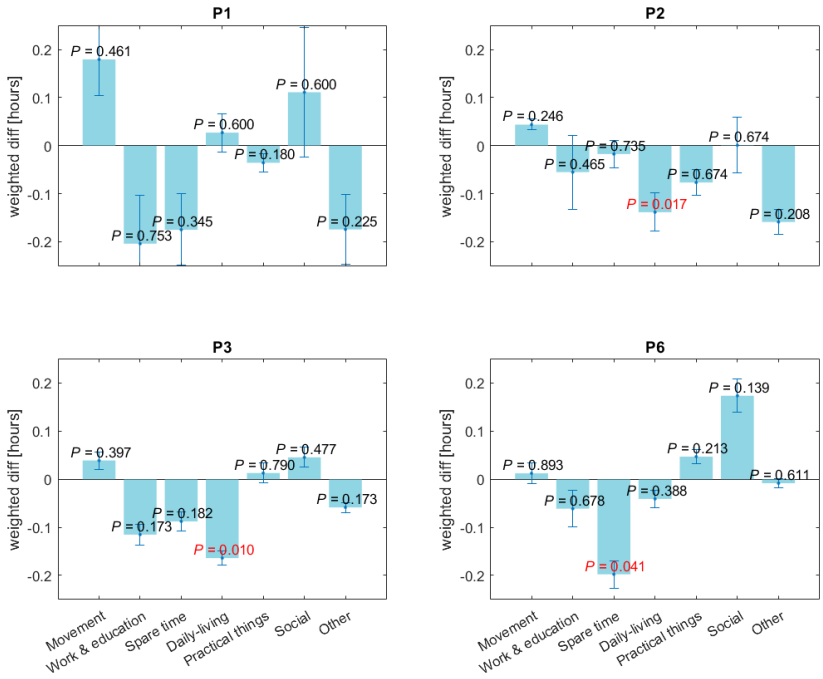


Figure 3.1: Plots from four patients. The bar charts show the fraction of hours of the seven AC. Positive values indicates a larger proportion of the AC during 'good days', with the inverse indicating more of the AC during 'worse days'. Results from the non-parametric Wilcoxon Signed-Rank test on the fractions is shown with the attached P-value. Red indicates a statistical significant result.

Wilcoxon Signed-Rank analysis, uncorrected for Multiple Comparisons (MC). We chose the Wilcoxon Signed-Rank as we were dealing with within-subjects (correlated samples) that did not follow a normal distribution, and we had two conditions ('good days' and 'worse days').

Common to all patients we see that 'Movement' (i.e., all types of physical activities) and 'Social' (i.e., activities involving other people) are positive (i.e., mostly represented in 'good days'), which are consistent with studies on physical [180] and social activity [71] for depressive symptoms. Interestingly, we see that the distributions are personalized to each patient. For example, P6 has an over-representation of 'Spare time' activities in 'worse days' (n = 44 hours), contrary to 'good days' (n = 27 hours), which was statistical significant ($Z(10) = -2.10, P = 0.04$).

Therapists provide general guidelines in line with our findings. It is always good to

schedule physical and social activities such as "*take a small walk around the local lake*" or "*call your mom and ask about her day*". However, as shown in the example, other activity patterns are highly personalized. Hence, the theoretical model on activity data suggests the feasibility for data-driven learning from self-tracked activities, which could equip both the patient and therapist with knowledge on personal behavioral traits and help the patient understand what activities impact mood.

3.2 Pooling correlation insights between sensor data and mood

Increase in battery capacity, computational power, digital distribution platforms, mobility, and embedded sensors has transformed smartphones from supporting voice communication to advance across many applications domains. Smartphones blend into everyday life and support social networking, education[2], and financial tasks (e.g., payments, and transfers). As such, smartphones have become a mediator for behavioral research. Particularly, affective disorders such as bipolar and depression are examples of symptomatologies that affect behavior. Figure 3.2 gives an example of the rich behavioral data that can be inferred from smartphones. GPS sensor data from a participant using `Moribus` during the summer of 2017, provides exact information on mobility, which can be cross-referred with the Foursquare API[3] for location information. In this case, the participant confirmed travel activities in respectively west of Jutland and the northern Zealand of Denmark.

Numerous studies exists to track, monitor, classify, and even predict depressive disorders from smartphone sensing (e.g., [170, 164, 60]). Notably, there are several cases of contradictory results. As an example, Faurholt-Jepsen et al. [65] found a statistically significant positive correlation between the number of outgoing SMS messages and depression, whereas Beiwinkel et al. [20] found a statistically significant negative correlation. Attempts, by systematic reviews, have been made to provide an overview of the different studies on mobile recorded subjective and objective features from users with affective disorders [52, 41, 158, 73]. Their focus is to summarize what exist, what has been done, and what are the barriers. It follows a qualitative study-by-study evaluation and does not try to answer what data relates most to depressive symptoms. We ran a large systematic review to provide a more quantitative comparisons of existing studies and find consensus on the most informative sensor data regarding correlates with depressive symptoms.

The systematic reviews is published and detailed in [A.2]. We searched 10 literature databases, resulting in $N = 46$ papers. We combined the studies to provide a detailed overview on the sensors, and methods used. For instance we identified 7 types of sensor data, what we refer to as Feature Category (FC) (Figure 3.3), 17

[2]https://kahoot.com/
[3]https://developer.foursquare.com/

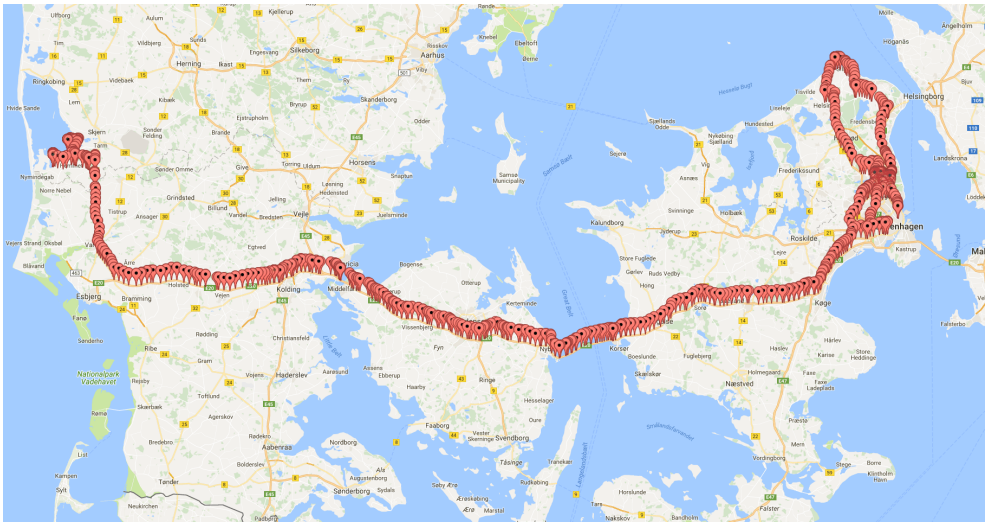


Figure 3.2: A map of Denmark with raw GPS coordinates as collected from a prior non-anonimized version of `Moribus`.

different sensor modalities (see Table 3.1 for the relation), and from these 85 unique objective features were recorded.

In the further analysis to investigate the correlation between behavioral objective features and depressive mood symptoms, we carefully split the analysis for research findings using non-clinical undiagnosed participants (n = 20, usually recruited from universities) and clinically diagnosed patients (n = 26, usually recruited from psychiatric clinics). The heterogeneity across the studies was too severe to perform any direct meta-analysis on the correlation values. In all aspects of a correlation analysis, the studies utilized different- (i) depression assessments, (ii) apparatus and frameworks impacting the measured features, and (iii) analytic methods. Therefore, to compare the correlation analysis, we derived a theoretical model (wD) that takes into account the correlation *directionality* weighted with the study sample size. The model definition is presented in details in [A.2].

Below we highlight some of the results for the clinical sample presented in [A.2]:

- Negative correlation between the number of cellphone tower ID's encountered and depressive symptoms
- Positive correlation between the duration of the smartphone screen being on and depressive symptoms
- Negative correlation between vigorous activity as measured by the accelerometer and depressive symptoms

	Subject	Social	Physical Activity	Location	Environm.	Device	Bio
FM radio signal				[124]			
Pedometer			[57]				
Camera	[194]					[8]	
Accelerometer	[47, 121, 155, 129, 178]		[21, 8, 19, 47, 60, 121, 124, 1, 20, 24, 62, 72, 75, 80, 84, 103, 112, 144, 148, 178, 182]				
Call log		[38, 56, 136, 128, 195, 20, 64, 63, 65, 80]					
Microphone	[150, 49, 61, 80, 81]	[150, 21, 23, 1, 136]					
SMS log		[194, 20, 64, 63, 65]					
Multiple-sensors	[21, 195]			[150]		[61]	
GPS			[195, 60, 163, 164, 20]	[21, 56, 195, 34, 40, 60, 163, 164, 1, 80]			
Bluetooth				[195]			
GSM				[20, 63, 65]			
Notification						[126]	
ECG			[62, 64]				[62, 64]
Wifi				[21]		[136, 128]	
smartphone screen						[136, 195, 8, 126, 128, 164, 3, 63, 65]	
App						[136, 126, 128, 3]	
Internet				[136, 128]	[56]		

Table 3.1: An overview of studies that contributed to the correlation between objective features and subjective assessment of depression or equivalent. The search was done in 2017, and is distributed in a 17×7 matrix with sensor modality and FC, respectively. Environment covers data such as humidity, temperature, and atmospheric pressure.

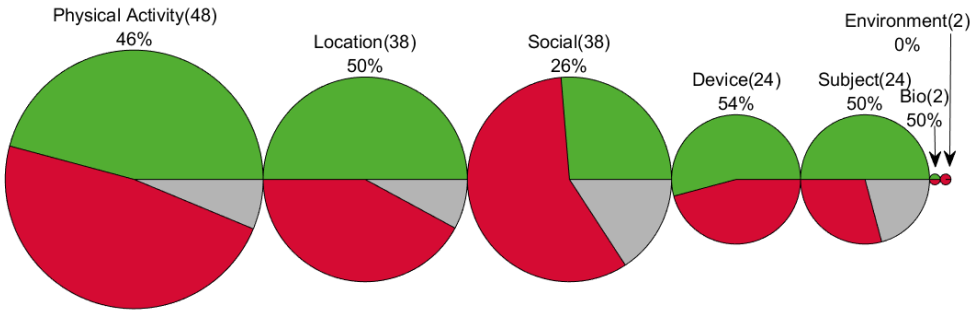


Figure 3.3: Each pie chart represents a category of features. The size of each pie chart indicates the number of different features included in the different studies. Green color reflects the number of features that statistically significantly correlated with depressive mood symptoms, red indicates statistically non-significant correlations, while grey indicates missing information.

- Indecisive correlation between number of SMS sent and depressive symptoms
- Indecisive correlation between the distance moved and depressive symptoms

3.3 Method for recommending personalized activities

Our theoretical papers in Section 3.1 and 3.2 focused on smartphones as a medium to self-track activities, and the recording of passive sensor data that can infer on the users current depression stage. The next step in our theoretical research was to investigate the capabilities of computerized sensemaking to provide actionable feedback. A smartphone is a pervasive technology that can be accessed to receive personalized and timely support, treatment and intervention when most needed. As smartphones enable collection of self-tracked- and passive sensed behavior, they allows for statistical and probabilistic techniques to automatically extract and learn from these data sources – what we showed in [A.1]. This is referred to as Machine Learning (ML) [94]. Three hundred studies, identified in a newer review paper, within mental health have used ML [171]. Based on these they conclude that "*ML demonstrate the potential to improve the efficiency of clinical and research processes and to generate new insights into mental health and well-being*" [171, pp. 10].

A majority of the studies are focused on advanced diagnosis and detection while only a few have designed a ML algorithm to provide intelligent feedback to the user. Hollis et al. [87, 177] used a linear regression algorithm to predict tomorrow's mood.

Inputs were historical mood data and the type of activity done. Therefore, they could estimate what type of activity to introduce in the model to shift the mood forecast in a positive direction. The estimated type(s) was then recommended to the user. The recommendations were, however, not *specific* activities. A core ingredient in BA is to schedule exact activities, as specificity significantly increases the probability of users enacting those activities, [173]. As presented in Section 2.2.3, the MOSS system [193] recommends activities. However, the recommendations were only based on past ratings, which does not follow state-of-the-art recommendation system methodology.

In our research, we have exploited state-of-the-art recommender research by developing a theoretical probabilistic multinomial NB classifier to model past activities as a bag-of-words[4] to predict mood outcome. The work has been published and presented in [A.3]. Briefly, we gathered activity, mood, and 'Pleasure' score data from two previous user studies containing 1,684 [B.1] and 3,921 [87] activity entries with corresponding mood and 'Pleasure' score from clinical- and non-clinical patients. The analysis was done offline, with the data split in a 'training' and 'test' set to estimate the performance of predicting what activities were enjoyable or otherwise. The split was done with incremental 'training' size to see how fast the model converges. As independent variables for the two models we used the specific activity text (A_1) and the corresponding labeled AC (A_2). The 'Pleasure' score was transformed to a binary dependent variable (C) indicating an enjoyed activity (C_1) or otherwise (C_2). We created a MATLAB (version R2018B) analysis pipeline to pre-processed the activities (e.g., removing stop-words, and porter stemming), as detailed in [A.3]. Subsequently, the activities were transformed to a bag-of-words (d_D). For the NB model we used the naive assumption[5] on the Bayes theorem, and hence our recommendation estimate was based on the conditional probability model seen in Equation 3.2.

$$P(C_j|D) = \frac{P(C_j)\prod_{i=1}^{F} P(A_i|C_j)}{P(A_1, A_2)} \tag{3.2}$$

The highest posterior probability for the two cases was selected as the outcome of the recommender model.

We compared our NB model with a SVM model, a generalized model where we pooled all participants, and a baseline model where we classified activities based on the most represented class[6]. To statistically compare the models we used a non-parametric permutation test with t_{max} based correction for MC [28]. We control for the MC problem as we are comparing the models on each sample (c_i), using a repeated measures t-statistic. The t-test for one of the samples (c_i) is shown in Equation 3.3.

$$t(N_{rep}-1) = \frac{\mu_s}{\sqrt{1/N_{rep}\sum_{j=1}^{N_{rep}}(c_i(j)-\mu_s)^2/\sqrt{N_{rep}}}}, \mu_s = \frac{1}{N_{rep}}\sum_{j=1}^{N_{rep}} c_i(j) \tag{3.3}$$

Several findings were uncovered from the clinical patients:

[4] The act of splitting activity entries into separate words

[5] that all our input features are mutually independent, conditioned on our outcome variable C

[6] Which is equivalent to the prior probability

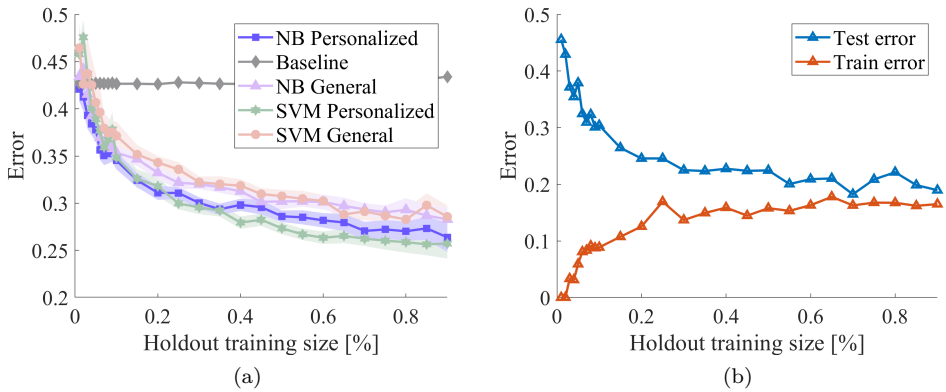


Figure 3.4: (a) The error rate as a function of training size of the different models for the clinical sample. SE is shown as a shaded interval. (b) training and test error is visualized for P5 on the NB model.

- The model converged fast dropping to 25% error rate (1 minus the fraction of test samples classified correctly to the total number of test samples) for the personalized version with the plateau situated after around 40% training sample size[7] (Figure 3.4). Figure 3.4b shows an example from P5. 40% training sample size corresponds to 100 activities.

- When comparing the personalized with the generalized model, we see a statistical significant difference after 15% training sample size (M = 45, SD = 16 activities) on both the NB and SVM model.

- Using a Kullback-Leibler divergence analysis on 90% training sample size to test the information gain of adding activity text or AC to the prior probability, we see contradicting results among the patients with some favoring activity text and vice versa. On a general level (for the NB model) the activity text ($D_{KL}(P(C|A_1)||P(C)) = 0.017$) carried more information than the AC ($D_{KL}(P(C|A_2)||P(C)) = 0.001$).

- By extracting knowledge from the trained ML model (NB model with 90% training sample size), we were able to conclude model-based insights: Figure 3.5a shows weightings of the AC feature, while Figure 3.5b shows a word cloud for the trained bag-of-words. The size represents stronger weightings. For instance we can argue that P1 favors 'Daily-living' based activities, and enjoy activities that includes words such as 'cosiness', 'radio', and 'university'.

[7] 40% of the data is used to train the model

- We incorporated a database of 320 pleasurable activities [119]. With the database we explored, through a simulation study, the effect of introducing novelty in the recommendation models. With training data from P1 and P4 we simulated the type of recommendations they received at three different time-points during the study.

Initially the generalized model performed best. This is equivalent to a real-world scenario where a recommender model simply does not have sufficient data to draw any reliable recommendations[8]. However, the error rate was swiftly reduced as more activities were used as a training sample. On average, after 58.92 (SD = 20.96) activities, the personalized model was favoured (M = 45, SD = 16 for the clinical patients only). We implemented the model in the `MUBS` system to test the performance and feasibility in a clinical study, as presented in [B.2].

3.4 Summary

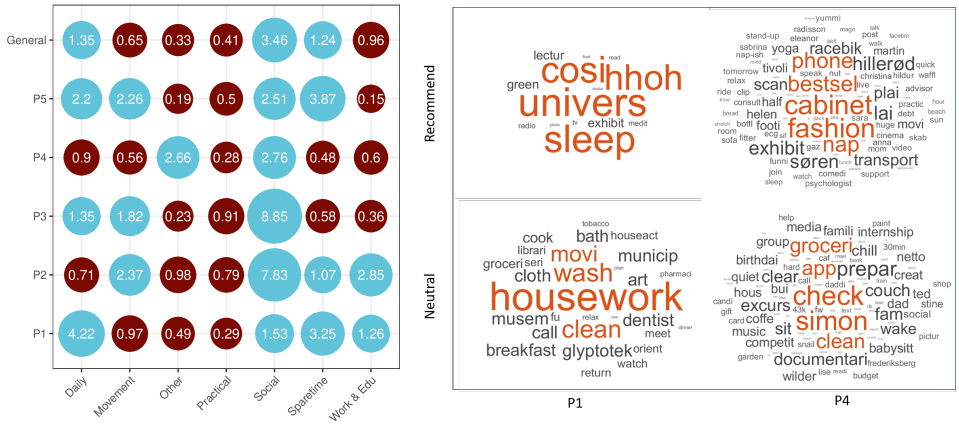
In this chapter, we presented three different theoretical models to leveraging patient generated data for BA support:

- A statistical model for analyzing patient self-reported activity data.
- A theoretical model for analyzing correlation between passive mobile sensors and depressive symptoms.
- An activity recommender model based on activity data sampled from mHealth technology.

These contributions demonstrate the feasibility of moving a significant portion of BA therapy to a mHealth system. First, through the potentials of statistical analysis on hand-written notes on behavior, the activity registration part can be supported. Then,by using digitally registered activities, the analysis was expanded to demonstrate the feasibility of supporting activity scheduling, which is then second part of BA. Finally, by modelling self-registered activities in a NB recommender classifier, we were able to predict activities as either enjoyable or neutral. Through a simulation study we demonstrated how the trained model could assist in the scheduling of activities through recommendations.

Our theoretical contribution in this chapter depicts the potential of statistical methods on BA data and sensor data. However, transforming them into a usable, useful, and clinical valid system, which is mediated in a suitable manner to a fragile target group, is another important consideration. In the next chapter we present the process of designing and implementing a suitable mHealth system that supports BA.

[8] In recommender terms this is also known as the cold-start problem



(a) (b)

Figure 3.5: NB weightings of the features trained on 90% sample size. (a) The ratio of the likelihood function for the AC. (b) Word cloud showing terms (after pre-processing) highly associated with recommended and non-recommended (neutral) activities for P1 and P4. The size of the unique words represents the fraction value. The words are presented after pre-processing and Porter stemming. Therefore, words such as 'race biking' has been transformed to its root form: 'racebik'.

CHAPTER 4

Designing mHealth systems

Coloplast[1] is the leading company for development and distribution of products for intimate needs - primarily Ostomy care[2]. In 2015 Forbes Magazine pointed out Coloplast to be the 4th most innovative company in Europe [68]. Years before this nomination however, the company spent millions of dollars during a two year period on the design of a product they had high hopes for. They redesigned the bag to be twice a durable. Yet, the entire project was discarded. They forgot to include ostomy patients during the design and missed the fundamental fact that the patients had no intention of wearing a bag, potentially with stool and urine, for more than a couple of days[3].

• • •

Understanding the design for a smartphone application for BA treatment of depressive symptoms, is one of the most relevant processes of building a system. Lack hereof, may even produce harmful apps [138]. iBipolar, as an example, advised people in the middle of a manic episode to drink hard liquor to help them to sleep [6]. To design and develop a quality app the designers have to bridge the connection with the patients and with the clinical providers within the designated mental treatment [185].

This chapter describes the design and development process of `Moribus` [B.1], and `MUBS` [B.2]. Although the chapter is outlined in chronological order of the processes, it is important to note that the standalone `MUBS` app was designed and developed after the feasibility study of `Moribus`.

4.1 Design of mHealth applications

mHealth technology provides interventions and/or support that is accessible, engaging, and potent. Although this belief is shared among many researchers, it is not

[1] https://www.coloplast.dk/
[2] A pouch situated outside the body to collect bodily waste from the large intestine
[3] A story by Coloplats's Vice President for Technology Hanne Everland

necessary true. Depending on the design choices different throughput and usage is seen [36]. In essence a successful design is characterized by good usability. A standard (ISO 9241-210:2019) defines usability as:

> The extent to which the system can be used by specified users to achieve specified goals with *effectiveness*, *efficiency*, and *satisfaction* in a context that is specified by us

Additionally good usability should ensure that the app interface has a low learning curve (*learnability*), and be memorable so the user can return to the app without the need to re-learn the interaction (*memorability*) [139]. Good usability is not an adequate term when designing a product. One has to consider the entire UX [32]. Houde and Hill defined the experience with a triangular model [90]:

Role What function does it serve? Is the technology useful to the user's life.

Look and feel What does the user look at, feels, and hears when using it? The concrete sensory experience.

Implementation How does it work? The technique and components used to make the technology perform its function(s)

It is a very dynamic, complex and subjective phenomenon. For instance the UX of snowboarding:

> "*depends upon the weight and material qualities of the board, the bindings and your boots, the snow conditions, the weather, the terrain, the temperature of air in your hair, your skill level, your current state of mind, the mood and expression of your companions*" [32, pp. 424]

Incorporating and understanding good usability and UX within mHealth - to achieve user satisfaction - requires the adaptation and use of HCI research. HCI studies the interaction between users and computers (e.g., software, apps or hardware) and, naturally, spans multiple disciplines [54].

Within mHealth systems for mental health and depression, HCI research has been emphasized as an important contribution [165]. Analysis of mHealth systems have shown that they suffer from over-engineering [152], low engagement [111], and poor adherence [7]. Problems that might be identified and addressed using HCI design methods. Such methods helps to (i) gather detailed clinical usage data, engagement and reports of UX, (ii) understand requirements and context, and how people interact with the technology, and (iii) explore the strengths and challenges associated with different design choices [165].

In this section we will describe our HCI design approach which resulted in high-fidelity prototypes for `Moribus` and `MUBS`. The subsequent technical implementation of these prototypes and the deployment to both the Google Play Store and iOS App Store is described in Chapter 5.

4.2 Ideation phase

There are several constraints that need to be considered when designing for mental health and depression. One is the regular access to end-user [53]. As a fragile user group we had to carefully plan out when and where to involve the patients in the design process. Therefore, prior to involving the user, we utilized our regular access to the clinic and therapists experienced in BA (see [C.1] for a description of the participating stakeholders). While introducing HCI research into the clinic was mentioned as a beneficiary, it also carries challenges when collaborating across a multidisciplinary team. Prior research describes the experiences as both (1) difficult due to the different technical backgrounds [137], such as clinicians concerns with the pitfalls of introducing technology into practice; and (ii) The difference in research background [29]. For instance end-users are considered the primary 'experts' by designers, while clinicians considers, themselves, as the 'experts'. With that in mind, we facilitated bi-weekly meetings with the clinic, which continued between and after patient involvement. See [B.1, B.2] for details. In short, we carefully explained the intended design process to make the clinicians aware of the motivations and methods used. Rather than a two-arm RCT study to assess a causal clinical outcome, we planned a feasibility study of the technology - to gain insight in different aspects of the technology interaction, and whether it works as intended.

4.2.1 Personas and Scenarios

To identify user goals and needs, we initiated a storytelling process. Based on a persona[4], we all prepared a scenario[5] for a future mHealth system. Ideas for the future mHealth system was based on prior work done in the team [16]. One of the presented scenarios is shown in Figure 4.1. The scenarios were shared and discussed, and we identified several noteworthy objectives for a smartphone system supporting BA:

- Some overall goal(s) need to be established by the patient. These should be a reflection of common activities that are capable during periods with no depressive symptoms. The phone should be able to asses and support in the progress of reaching the goal(s).
- *Follow me, and not your condition* - The phone should provide clear agendas or plans that can be followed with least amount of effort.
- The phone should support the activation of following general recommendations: (1) **Sleep** - get up from bed early; (2) **Hygiene** - take a shower; (3) **Food** - have a nutritious diet; (4) **Outside** - get out the door; and (5) **Social** - establish some social contact.

[4] A fictive character representing a personification of a real user

[5] A mini story that reflects a situation the persona might be in, and how a possible technology would assist

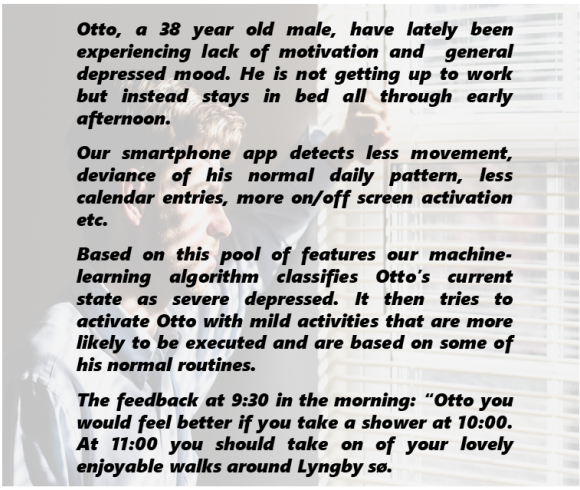


Figure 4.1: Scenario from design meeting 25.10.2016.

Towards the design of a standalone mHealth system we had to understand the role of technology in BA therapy. Our initial theoretical work [A.1] showed positive evidence on the feasibility of designing a digital platform supporting BA. Not only was it shown to be realizable based on a high activity registration compliance, but the statistical analysis provided additional insights than regular face-to-face consultations. With the suggesting that a digital solution could assist the psychologist in BA therapy sessions, we sought to design a smartphone app to support BA in therapy. Based on these findings the next step would be the design of a standalone smartphone system capable of inferring on the collected data to recommend exact activities to enact on.

4.2.2 Low-fi prototyping

Through an iterative design-process with the interdisciplinary team, we sketched the initial design proposals on paper-based mockups. For `Moribus` we digitalized the BA paper-based weekly schedule [A.1, Figure 1]. This is referred to as the 'Calendar' function of the app and is the core part of the BA method.

The first low-fi prototypes of `MUBS` were created based on insights from the design and feasibility study of `Moribus` [B.1]. This includes the change of an hourly activity assessments to a more adaptive *morning*, *afternoon*, *evening* structure. A structure that resembles classical to-do apps with a list of daily tasks. As `MUBS` was intended for users outside of the clinic, and not in collaboration with a therapist, it was streamlined

to the general population, with terminology change from depression to well-being, and symptom management to lifestyle management. MUBS had to bypass the stigma usually present when developing for a specific clinical target group. A great example of this is the peeler example by Norman [142, pp. 244]. Furthermore, opposite to Moribus, a goal was needed, and initial design focused on the balance between the five general behavioral guidelines listed in previous section. See Figure 4.4 for three early sketches of MUBS.

A central part of HCI design principles is the involvement of the end-user in a UCD process. It was a strong asset in the design of both of our applications. We used UCD in an iterative process where we used the user inputs to refine our design choices and re-scheduled new meetings for new rounds of insights.

4.3 User-centered design

UCD comes from HCI and was originally a software design method for developers and designers but is applicable for all kinds of product developments. The goal of UCD is to ensure that the users needs are met, that the resulting product is understandable and usable, it fulfill the desired task, and the experience of use is positive and enjoyable [142]. UCD methods ensures that the focus rests with the end-user, the people that are eventually going to be the core users of the product, by involving them in the design process. The level of involvement can range from minimal contact where clarifying questions are unraveled by a given point in the process, or maximal contact throughout the design process. The latter was used in the development of our two apps where the users where involved in the initial design phase but also during prototyping and in subsequent field studies, evaluation and interviews. In the initial design phase we invited patients to full-day workshops where we began an early evaluation of design concepts, and targeted the five usability heuristics mentioned in Section 4.1.

4.3.1 Workshop

We held two hours workshop days in the design phase for each app. See the method section of [B.1, B.2] for details. Figure 4.2 shows pictures from some of the workshops. We carefully planned out a workshop agenda, and assigned a facilitator. See Appendix E for the protocol of part 1 of one of the MUBS workshops. As suggested by both Doherty et al. [53] and Bardram & Frost [14] we followed the Patient-Clinician-Designer Framework [120] to structure the discussions and considerations during the workshops. The framework includes the needs and perspectives of clinicians and designers thereby mediating co-design with patients. Lets take an example where a patient expresses a desire to include a chat-room functionality in the system. The patient wants suggestions from fellow patients on actionable decisions. The PCD framework then includes the clinicians concern of acceptance by asking: "*Does the*

system encourage patients to make decisions without consulting a clinician?". Although the aim is not a positive answer to that question, it assists in reaching a shared consensus. Together they might guide the design towards chat-rooms that are only open at specific times when a clinicians is online and can participate.

`Moribus` Workshops We recruited three affective disorder patients; two bipolar and one major depressive disorder. Together with the designers and clinicians we were seven participants including one facilitator. The main goal of the workshops was to;

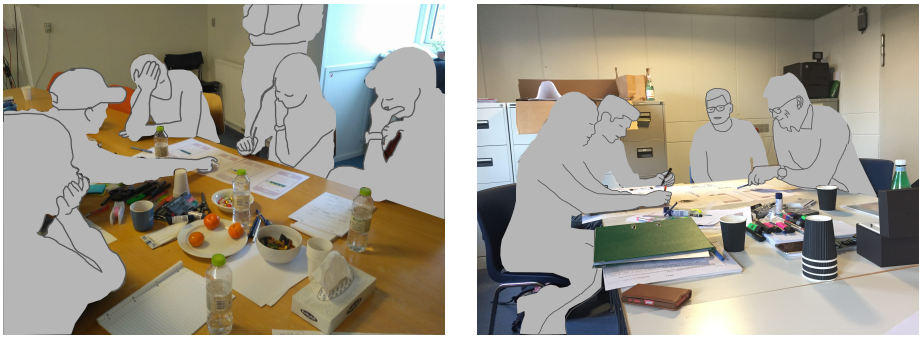
> gain insights on the design of hourly activity planning and registration on a smartphone

The room was decorated with example data visualizations (i.e., screenshots of apps within healthcare, such as the SleepTight app [39, Figure 3]), tangible hardware (i.e., smartwatches, and smartphones with different form factors), and inspiration ideas of interactive opportunities (i.e., screenshots of the phone unlock journaling where one can rate mood from the lock-screen [200, Figure 1]. Furthermore, we provided materials for stimulating the design of various different low-fi prototypes such as A4-sized papers with empty app screens, color pens, pencils, large poster papers, scissors, and tape. We initiated by an introductive presentation of the general goals of the app from both technical and clinical perspectives. Then we presented the design goal of activity scheduling and prior work in the field. This was followed by a brainstorm session where all participants used post-it notes to write three challenges and positive things about activity schedules. These were discussed, which led to the next task of hands-on low-fi prototyping. We split in two groups, and drew out ideas of a smartphone app with activity registration. We used all the available materials to cut and glue prior ideas and new mockups together (see Figure 4.3 for some examples). All mockups were gathered together on the large poster papers, and presented between the groups. Ideas and feedback on the next workshop were used to progress from sketches to wireframes to interactive prototypes.

`MUBS` Workshops We recruited four new patients; three with bipolar and one with major depressive disorder, and a new designer. In total we were 12 participants. The main goal of the workshops was to;

> gain insights on the design for registering daily tasks, the feedback methods, and the role of the recommender model

Similarly to the first workshop, the room was decorated and filled with inspirational material. In addition, images and descriptions of the different personas we places on the wall of the room (Appendix E, Figure 1), and we had music running in the background on low volume. Prior to the workshop we received demographic information on the participating patients, and carefully organized the workshop and the



(a) (b)

Figure 4.2: Two images from two different workshops. The participants have been anonymized through a silhouette technique.

additional modification, i.e., music, as one of the patients was a musician. We initiate with an introduction of the overall goal of the app, theoretical knowledge of BA, and lessons learned from the design and development of `Moribus`. The workshop was facilitated with to main activities. The participants were split in three groups but reshuffled between the two main activities. First activity was an active ideation phase by the participants. They were (i) told to create visualizations that could motivate registrations, and (ii) sketch interaction ideas on how users of the app system are suppose to register completed tasks. Second activity was a discussion and debate on three existing low-fi prototypes developed during the multidisciplinary team meetings (Section 4.2.2). The second workshop ran in three parts, (i) group discussions and brainstorm on new low-fi sketches (see Figure 4.4) focused within two areas, visualizations of daily tasks, and the recommended activities. (ii) An individual *mind-mapping* session to create a dialog within concrete questions. The questions had suggested answers, and were put up across the room. The participants had to vote for the favorite suggestion for each question, and could add comments and ideas by attaching colored post-it notes to the questions. (iii) With inspiration from the mind-mapping session, we ended the workshop with an idea sharing session. In pairs the participants sketched different visualizations of historical data, and how the visualizations could mediate a sense of progress.

Implication of the patient group Throughout the workshops we were observant of the different symptoms expressed by the patients. Some patient were more dominating, talking faster and for prolong period, while others exhibited more depressive symptoms and did not actively engage in the larger debates. As we were well aware of these behavioral scenarios we deliberately chose to do smaller group session and make informal one-on-on conversations in an attempt to receive inputs from all patients.

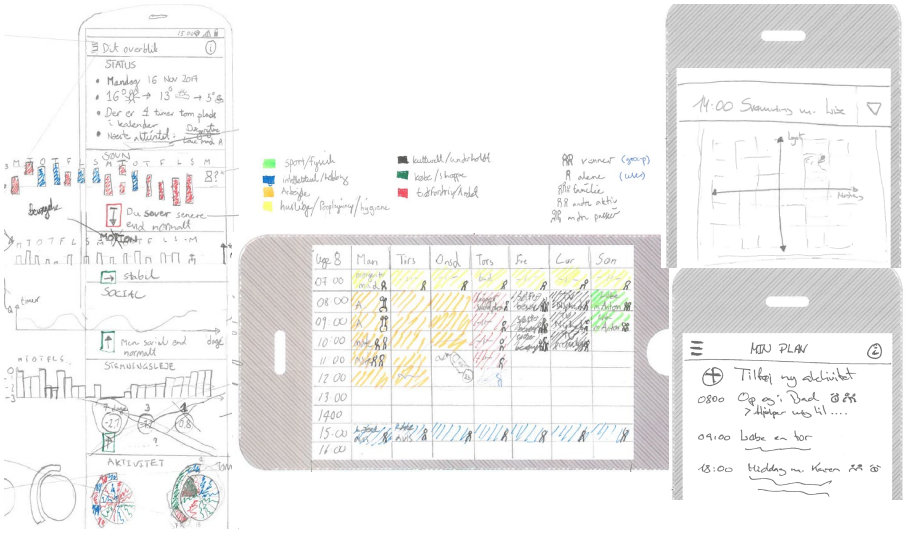


Figure 4.3: Design sketches `Moribus`. From left, an overview idea of self-assessment graphs covering sleep, physical activity, social activity, mood, and types of activities done. Then an early version of the 'Calendar' labeled with the AC and social context. In the top right, a way to register 'Pleasure' and 'Mastery' from the notification. Lower right, an idea of how to present the daily plan.

Figure 4.4: Design sketches `MUBS`. Illustrating on various way to present the daily plan. The two left operates on the idea to have a 'daily goals' that is accomplished when the star is colored, and all the leaves are marked, respectively.

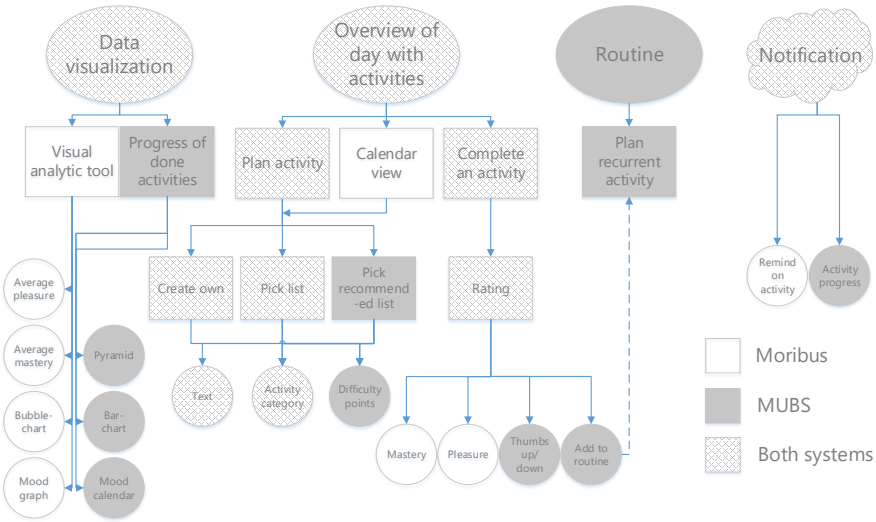


Figure 4.5: A flowchart showing some of the main functions in the systems. The dark grey illustrates features that are only implemented in `MUBS`. This includes the Routine functionality, and the recommender model. While the white is features only in `Moribus`, such as the 'Calendar' function.

4.3.2 Prototype iteration

In the interdisciplinary team we utilized the feedback and insights derived from the UCD activities to develop digital mockups, what we refer to as high-fidelity prototypes. See [C.1, Figure 2] for some examples on `Moribus`. The prototypes were developed in Justinmind[6] (vs 8.2.1). With drag-and-drop of UI components, and simple If This Then That (IFTTT) conditioned statements we were able to produce interactive prototypes. Through an iterative design process we finalized the design of `Moribus` and `MUBS`. The final design and the features are explained in [B.1, B.2], and shown as a flow chart in Figure 4.5. Functionalities implemented in both systems are illustrated with a patterned background. Table 4.1 provide an overview and description of the features implemented in each system.

[6] `https://www.justinmind.com/`

Feature	Description	Relevance for patient	Relevance for clinic
Activity Category	Six distinct classes covering all types of activities, as inspired from [145] and terminology within mental health and BA [135]	Able to quickly register an activity without writing text	Get information on what type of activities the patient usually indulge in
Pick list	A catalog of 384 (only 60 in `Moribus`) pleasurable activities by work of Lewinsohn & MacPhillamy [119] and Mørch and Rosenberg [135]	A source of inspiration to plan enjoyable activities. In `Moribus` it served as examples to understand the different categories	With less patient contact they want to be able to provide the patient with generally recommendable activities
Difficulty points	In a range between 1-3 it defines the effort it takes to perform an activity [B.2]	To be able to prioritize easier activities during period with depressive symptoms and vice versa	To understand the perceived effort of the patient to enact on basic activities
Visual analytic tool	A page in the app that provides summary statistics of the registered activities [B.1]	Gain data-driven insights. To explore the relation between activities, and mood and/or 'Pleasure' rating to improve self-awareness on behavior	To assist in their understanding of the patient, and find more complex relation between activities and depressive symptoms
Pyramid	A visualization of the progress in all six activity categories[B.2]	To discover what type of activities that have been neglected	To facilitate the feeling that a balance within the different life areas is important to get better
Recommendation	An implemented NB model [A.3] to recommend 10 personalized activities. Four of the activities have been done before, while six are novel. Novel activities are drawn from the 384 item catalog	To get inspired by activities that the system intended for the patient	With activities that are relevant for the given user, it could improve motivation to plan and initiate the activity

Table 4.1: An overview of the core design features in `Moribus` and `MUBS`.

4.4 Summary

This chapter provided an overview of the UX and HCI design methods applied in the design of the two mobile apps; `Moribus` and `MUBS`. We followed a UCD process to maximize usability of our target user population. Through a total of 4 workshops across the two systems, we discovered the importance of a simple UI, and the utilization of a point system to motivate the initiation of more difficult tasks. Patients were in general favourably disposed towards the use of these technologies, and expressed an desire for sensemaking between mood, 'Pleasure' and types of activities done. We also found that users were not fond of (i) recommendations being 'pushed' to their list of daily tasks, or (ii) daily goals with a pre-specified requirements of the number of activities that needs to be completed. Instead, they preferred a more passive – manual selection – design choice.

CHAPTER 5

Technology for Depression

In 1965, Engineer Gordon Moore observed that microchip capacity seemed to double every 18-24 months, equivalent to a fast increase of computational processing power. The observation became known as Moore's law and is to some degree still observed today[1]. Similar scaling is seen in software- capabilities, architecture and applications [96]. The fast paced technological progress carries severe implications for research. It enables HCI studies of exciting new technologies such as smartwatches [30] and virtual reality [10], but it 'closes' the relevance of others, as technology is quickly outdated. Although recent studies have shown that depressive symptoms can be inferred from the amount of SMS sent [A.2], it might be unimportant due to the significant drop in SMS usage [2]. The choice between a large scale RCT study spanning several years, versus a single-arm feasibility study spanning weeks might be crucial in the choice of study design for the interaction of technology.

• • •

We experienced the technological progress in the implementation between `Moribus` and `MUBS`. As a consequence, the one year older `MUBS` system was built with a different configuration than `Moribus`, even though it served as a natural next-version system.

5.1 Moribus

The system serves two purposes. From a practical point of view, to target the increased treatment gap, `Moribus` shifted parts of BA therapy to a digital smartphone system. It provides hourly activity planning and registration with access to visual analytic tools. These tools displayed statistical relationships between the users activities and corresponding mood, 'Pleasure' and 'Mastery'. As such, `Moribus` facilitated sensemaking and insights into one's behavior. Insights that served the purpose of improving the patients understanding of their depressive symptoms, and support the

[1] https://medium.com/predict/moores-law-is-alive-and-well-adc010ea7a63

clinicians during therapy sessions. Below we demonstrate a little scenario of the intentions behind `Moribus`:

1. During initial therapy, the clinicians install `Moribus` on the patient's phone, and tell them to user the app for (i) planning activities a day ahead, (ii) register what you did every hour, and (iii) register your daily mood.

2. The patient uses the app for two weeks and access the visual analytic tools in an effort to understand relations between behavior and symptoms.

3. The patient meets with the clinician. Together, with the data from `Moribus`, they locate activities that were positive or negative reinforcers, and discuss future behavioral goals. Based on the discuss they use `Moribus` to plan the following two weeks.

4. With the recommendation of the clinician to "*shut down your brain, and just follow the plan*" the patient uses `Moribus` and starts feeling the impact of digital aided BA therapy.

From a theoretical point of view, we were interested in building a database of behavioral data. Data on completed activities and their corresponding rating, and context. No similar database exists, but was needed to build an advanced algorithm to detect positive and relevant activities personalized for each user. An algorithm that was added to a next-version of `Moribus` presented in Section 5.2.

5.1.1 User Interface

As explained in details in paper [B.1], the patients are presented with an hourly schedule as they enter the app. The schedule is shown with a pie-distribution of the types of activities as categories by AC. Through a navigation menu the patient can enter other functionalities (e.g., the visual analytic tools). Screenshots from the app are visualized in Figure 5.1.

5.1.2 Software architecture

An overview of the software architecture is provided in Figure 5.2. We handled the passive sensor data, and the subjective recorded activity together with mood data separately. The sensor data was collected into a JSON file and securely transmitted to the Amazon Web Services (AWS) Simple Storage Service (S3). As we carefully explain in [B.1], we do not store any patient identifiable data. Instead the patient is represented by a 4-digit PIN. Figure 5.3 shows the file structure within AWS.

The subjective data was stored in a local SQLite database. This gives us the opportunity to read and write to the database independently of network data and utilize SQL queries for the visual feedback tools. A quick overview is provided in

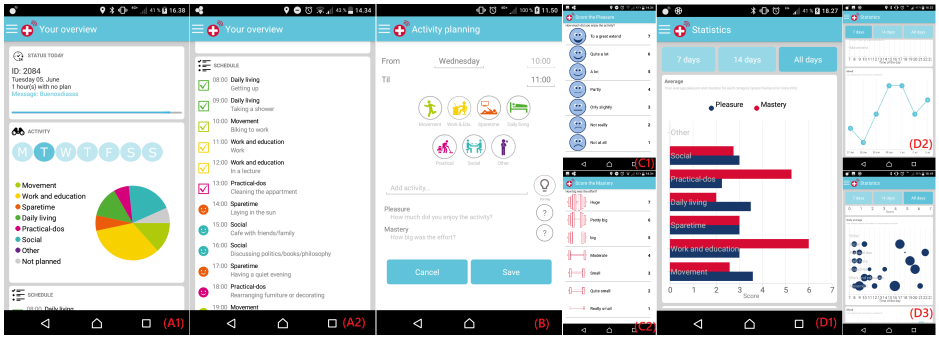


Figure 5.1: The `Moribus` mobile user interface. The home screen (A), shows a pie chart summarizing planned and done activities within each category (A1), and a list of planned or registered activities (A2). (B) shows the page for creating an activity, while (C) illustrates the pages for rating an activity with 'Pleasure' (C1), and 'Mastery' (C2). The visual analytic page (D) summarizes historical completed data. The bar chart displays the average score of 'Mastery' (red) and 'Pleasure' (blue) for each activity category. The bubble chart shows the average 'Pleasure' score for each activity category across time. The line graph displays the reported mood scores over time.

Figure 5.4. It illustrates the origin of each field in the activity database and how it is used to display the bar chart in the visual analytic tool.

5.1.3 Implementation

We searched existing mobile sensing toolkits to fulfill the requirements of sampling sensor data. Particularly, we were interested in (i) a toolkit supporting both Android and iOS, (ii) up-to-date maintenance, (iii) support of GPS, and device based data sampling, and (iv) preferable written as a cross-platform framework in a single programming language. A larger analysis of the different frameworks was out of scope for the dissertation. Instead we provide an overview of the different toolkits in Appendix F.

We chose Sensus [197] because it integrates well with a cross-platform developing toolkit named Xamarin, which is based on .NET technology [154]. Working with a cross-platform developing toolkit has two main advantageous. First, all code was written within the same development environment. Second, even though the app is native on both Android and iOS, the code is almost identical for both platforms. In this way, developing apps in both Android and iOS was possible in a much shorter time frame and within a smaller developing team.

We used the Shared Asset Project (SAP) method in Xamarin. This implied that

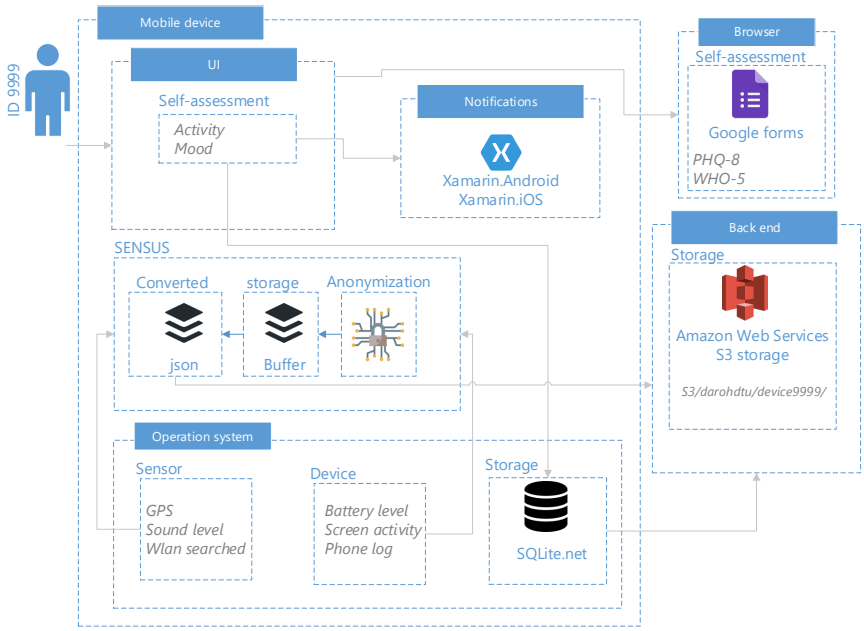


Figure 5.2: The software architecture of `Moribus`. In the upper level we have the mobile device, and a browser link to a Google Forms questionnaire on PHQ-8 and WHO-5.

the shared- code and assets were used in both the iOS (Xamarin.iOS) and Android (Xamarin.Android) .Net libraries during compilation. The Xamarin library features several UI classes. Lets take the case of a `switch` component. During compilation the Xamarin C# compiler would generate an Intermediate Language (IL). Here, in the individual renders, the `switch` is mapped to a `UISwitch` on iOS and `Switch` on Android. For iOS, the compiler then uses the Apple compiler on a Mac to generate native iOS machine code (like the Obj-C compiler), while on Android the IL runs the Mono runtime framework that uses the Java engine to generate native Android code [154]. In this way the shared code base manages to produce native looking UI components in the app. To accomplish an overall native and familiar experience to Android and iOS users we furthermore addressed the design guidelines defined by Apple[2] and Android[3].

Unlike typical frameworks, Sensus was configured as an entire system. Researchers, without any programming requirements, could configure so-called sensing plans. In

[2] `https://developer.apple.com/design/human-interface-guidelines/ios/`
[3] `https://material.io/design/`

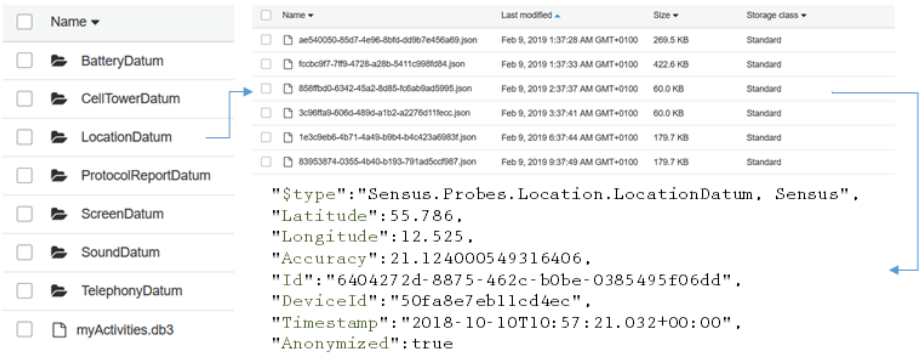


Figure 5.3: The data structure of the stored sensor data in AWS. This example shows the location folder with its list of JSON files.

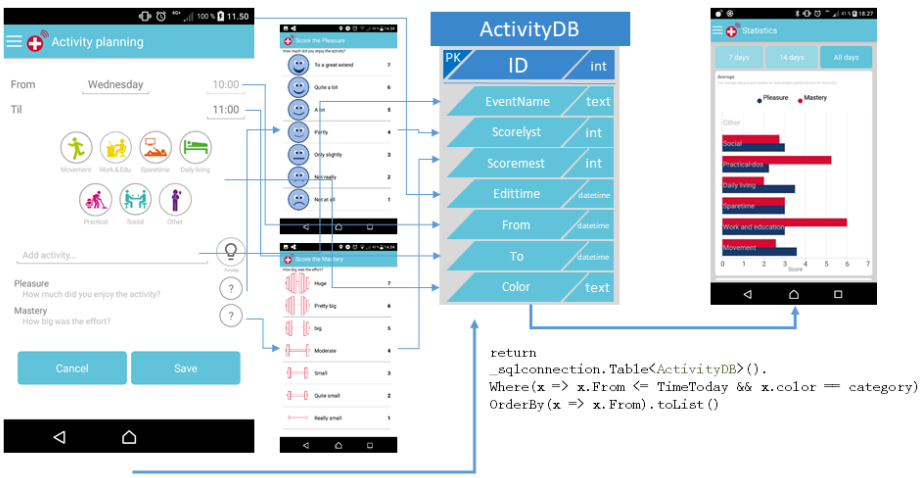


Figure 5.4: An illustration how activity data is stored in `Moribus`, and how it is used to show the bar chart.

the sensing plan the researcher has the ability to choose what sensors to sample from and schedule surveys. Subsequently they can distribute the plan to the users. As we were only interested in the sensing framework, we followed an bottom-up approach. We downloaded the source code of Sensus and initiated the implementation on top of the Sensus sensing framework, discarding the unnecessary UI components and sensing plan configurations. To most efficiently share as much code as possible we followed a Model-View-ViewModel (MVVM) architecture where we separated the code in a Model (carrying data and fixed values), a View (the UI components), and a

ViewModel (the interface between the Model and the View). The code can be found in the CACHET GitHub repository[4].

The implementation of `Moribus` was successfully published to the Google Play store[5] and iOS app store[6] in August 2017.

5.2 MUBS

The goal of `MUBS` was to develop an app that - independent from the clinic - supports users experiencing depressive symptoms. The term 'independent' was used to emphasize that the app automatically, without the assistant of the clinic, provided actionable feedback personalized to the user. The feedback was designed as an activity recommender. Similar to Step 4 in the `Moribus` scenario, the recommender was able to classify and propose positive activities to assist in the planning of enjoyable future activities. A typical scenario of the `MUBS` is outline as:

1. In the evening the user opens `MUBS` and plans a task for the next day. They have only one particular task for tomorrow: "*Eating dinner with my mom*".
2. Their planned day looks empty as the planning page is organized within morning, afternoon, and evening. Therefore, the user is interested to schedule an activity for the morning. In the recommender page, the user inputs tomorrows date, and specifying a morning activity.
3. Ten different activity recommendations is instantly shown in `MUBS`. Due to their lower energy level, they pick a recommended activity with a difficulty score of one: "*Make a cup of coffee and read two articles in my favorite newspaper*".
4. Next morning the user wakes up and opens `MUBS`. They inspect the plan and follow their scheduled morning.
5. Before bedtime the user registers today's planned activities as "Done" and rates them both with a thumbs up. Afterwards, the user inspects the pyramid visualization and see that there has been less focus on 'Practical' activities. As a consequence the user enters the recommender page, requests an afternoon activity and filters to only show 'Practical' activities. the user picks one of the activities and add it to their schedule.

5.2.1 User Interface

Screenshots for the implemented application are visualized in Figure 5.5. The bottom tab bar features four menus as explained in [B.2]. Briefly, we provide a summary of the

[4] `https://github.com/cph-cachet/radmis.Moribus.app`
[5] Was removed on October 2018 due to new policies on phone log sensing
[6] `https://apps.apple.com/us/app/moribus-adf%C3%A6rdsaktivering/id1271348746`

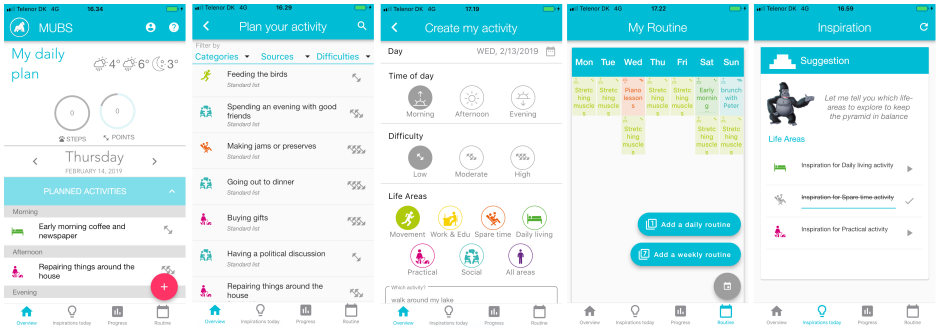


Figure 5.5: The MUBS mobile user interface. From left to right: The main screen with an overview of planned and done activities. The planning page where you can search or create your own. The page that is displayed when you create your own activity. The routine tab, where the user can create reoccurring activities on a weekly basis. The inspiration page with three activities within activity categories that the user is recommended to focus on. Pressing one of them gives a list of ten recommended activities within the selected category.

different menus. 'Overview', has information for the selected day. The activities are shown in a list separated in planned and done activities, which is further separated in morning, afternoon and evening. With the red plus sign, the user is able to either search a catalog of activities (Figure 5.5 2nd) or select a recommended activity. The recommended activity is calculated by the NB machine-learning algorithm as described in [A.3].

In the activity catalog, when the user searches for an activity that does not exist, they get the option to create their own (Figure 5.5 3rd). The 'Inspiration' menu features three suggests categories of activities for the day (see Listing 5.1). When pressed, it redirects the user to the recommended activities filtered by the chosen category. The 'Insights' menu displays past days registrations as a calendar. The days are colored according to the registered mood. A recap of the registered activities is shown when a day in the calendar is selected. The 'Routine' tab is a way to plan recurrent activities. A routine activity will automatically get added to the planned activities for the respective days.

5.2.2 Software architecture

An overview of the software architecture is provided in Figure 5.6. Xamarin has been replaced with the a cross-platform developing toolkit called Flutter[7], developed by Google. It was in an early beta-stage during the development of MUBS. As it was

[7] `https://flutter.dev/`

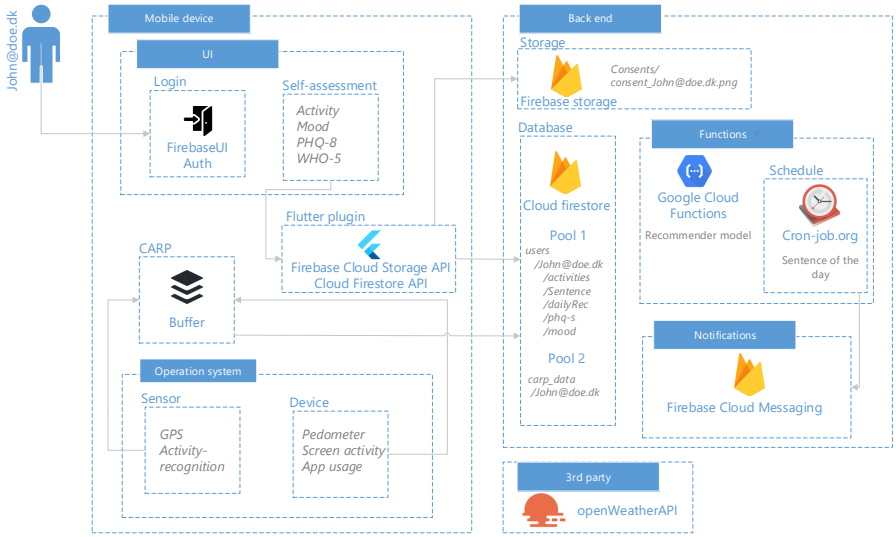


Figure 5.6: The software architecture of MUBS. In the upper level we have the mobile device, the backend, and 3rd party integration.

developed by Google it had well documented integration with the backend service Firebase[8], and the Google Cloud Platform (GCP) to host and execute functions. The latter was the main reason to use Flutter, as GCP could host the recommender model [A.3]. To record data from the phone sensors, we used our newly developed framework that supports Flutter and data storage in Firebase. The framework is called CARP, and is now publicly available in the dart repository [13]. In reference to Appendix F, it supports the three displayed features. We chose to remove phone log sensing as Google Play changed their policy as of October 2018[9]. Google does not allow the sensing of phone logs (calls and sms) unless its a core functionality of the app. The recorded sensor and device probes are shown in Figure 5.6. Every sampled data point is streamed to a Firebase database. The database structure for the sampled probes is provided in the GitHub page of CARP[10].

In Table 5.1 we provide an overview of the externally implemented functions within

[8] https://firebase.google.com/

[9] https://android-developers.googleblog.com/2018/10/providing-safe-and-secure-experience.html

[10] https://github.com/cph-cachet/carp.sensing-flutter/wiki/B.-Sampling-Data-Formats

the GCP. The functions are directly linked to the Firebase database and the mobile app, and offers a way to develop app functionality that does not compromise the phone's resources. Three functions covered by `calcDailyRec`, `itemRecommend`, and `dailySentence` are elaborated on as they constitute a significant part of the app.

`calcDailyRec` is run every night. It queries all done activities for the past 7 days and calculates the average difficulty points gathered for each AC. Listing 5.1 displays a snippet from the function.

The calculated average (line 23) is compared to a average of the past six days where the seventh day ('today') has values of 0 (line 24). The three categories that have the highest negative deviation (line 29) will be returned as suggestions in the inspiration tab menu. As an example, if the user does a single physical activity (e.g., swimming) of 3 difficulty points every Wednesday, s/he will have a deviation of zero up to 6 days after. On the next Wednesday the deviation is -3 and will be displayed as a suggested activity for that day. The average values are compared to indexed values in levels of 1-3 (see Figure 5.7) to estimate the placement on the pyramid – our analogy of user progress. The index values are derived from 4 weeks of gathered activity data in the `Moribus` field study [B.1]. It was based on a linear combination between the average rating of 'Mastery' for an AC and the amount of activity within the AC, where a low 'Mastery' together with a large quantity of the AC corresponded to a high index value. As an example, if you do a 'Work & Edu' type activity with a difficulty level 1 everyday (index = 7/7 = 1), you will be in level 2 on the pyramid (1.39>1>0.93).

`itemRecommend` uses the NB algorithm presented in [A.3]. We combined the likelihood function of words (as illustrated in Figure 3.5b) with the likelihood function for

Name	Trigger	Description
`sendWelcomeEmail`	User created	A welcome message is sent to the user's email
sendByeEmail	User deleted	A confirm message is given to a deleted user
sendConsentEmail	Consent approved	A copy of the consent form is sent
createMood	User created	A 'submit your mood' activity is added to the planned activities
createRoutine	'rout' = *True* in `activity`	The activity is replicated 8 weeks ahead of time
itemRecommend	Request recommendation	Calculates the probability of 'thumbs up' on all activities from the catalog (that have not been done before) and all previously done activities, and submits the top 10 activities to the app
weeklyPHQ[a]	Daily 10 am	Checks whether 14 days have passed. If so, creates a planned activity to fill out PHQ and WHO-5
calcDailyRec[a]	Daily 03 am	Calculates daily suggested AC to plan out
dailySentence[a]	Daily 07 am	Based on the done activities, the function calculates whether you increased in 'level'. Based on the 'level', a sentence is submitted as a notification message.

[a] Added as a daily cron-job triggered by `https://cron-job.org`

Table 5.1: An overview of the functions hosted on GCP.

```
const activityRef = db.collection(`users/${userRecord.email}/activities/`);
var queryRef = activityRef.where('done', '==', true);

// Calculate for each category:
outResult = queryRef.get().then((snapshot) => {
    snapshot.forEach((item) => {
    // For yesterday (because now its 3 a.m.):
    var inputDate = new Date(item.get('day'));
    // Gather the points to calculate the current MA from the past week (7
        days) later on:
    if (inputDate.getTime() <= todayD.getTime() && inputDate.getTime() >=
        sevenD.getTime()) {
        var timeDiff = Math.abs(inputDate.getTime() - todayD.getTime());
        var diffDays = Math.ceil(timeDiff / (1000 * 3600 * 24));
        movingAvgRaw[diffDays][item.get('ac')] += item.get('flower') + 1;
    }
    });

    for (j = 0; j < 6; j++) {
        addNew = movingAvgRaw[0][j]; // 0 represents diffDays, which is
            difference of days, and present time is ofcourse 0.
        addOld = movingAvgRaw[7][j];
        for (i = 1; i < 7; i++) {
            sum += movingAvgRaw[i][j];
        }
        movingAverageOld[j] = (sum + addOld) / 7;
        movingAverageNew[j] = (sum + addNew) / 7;
        sum = 0;
    }
    for (i = 0; i < 6; i++) {
        // Difference based on Old vs New MA:
        dangerMA[i] = movingAverageNew[i] - movingAverageOld[i];
    }
}
```

Listing 5.1: Moving Average calculation.

the AC (an example shown in Figure 3.5a). The combined likelihood was then post-weighted with the likelihood function conditioned on the current context: time of day (morning, afternoon, or evening), and the day (weekday, or weekend). This made it possible to derive the posterior probability on whether the specific user will rate the activity with a thumbs up or thumbs down of any activity. We chose to present the user with the four activities with highest posterior probability previously completed, along with top six activities from the catalog of 384 activities that have not been done before. We drew inspiration from the EmotiCal app that has five historical and five new activity types [87, 177]. In this way we accommodate two scenarios. A case of familiarization, and a case of novelty.

`dailySentence` is hosted as a cron-job by `https://cron-job.org`. It runs every morning at 07 a.m. to generate a personalized morning notification. The generated

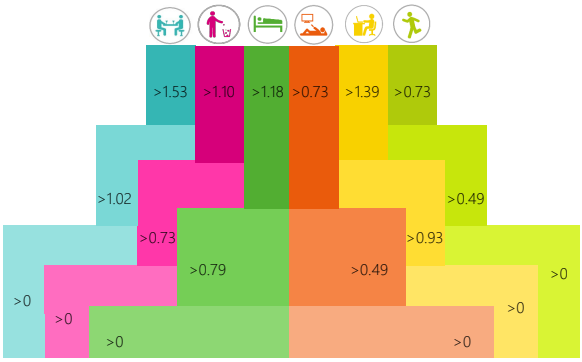


Figure 5.7: A breakdown of the different average valued difficulty points necessary in each level. The cut-off values are illustrated on top of their designated stair-case level on the pyramid. Level 1 and 2 have been shaded in lighter colors to show the subdivisions.

sentence is based on a tier-system. The lowest tier (tier-5) is during in-activity, while the higher tiers are are achieved when there has been an improvement in levels from one day to the next for any of the AC. The tier-system is stored in Firebase Storage. An overview of the tier-system is given in Table 5.2. If there was no improvement, we first checked if there was any done activities (tier 4). If not, we randomly selected a tier 5 message as the morning notification. If there was an improvement, a top-down approach checked if the user was in the top level of all AC's. If not, we went to the requirements for tier 1 and so forth.

5.2.3 Implementation

Developing in Flutter was done in the Dart programming language, while functions in GCP were written in JavaScript. With the Flutter framework all of the code base was shared between Android and iOS. Flutter uses its own rendering engine, and the code is compiled to native code.

The entire UI is built by individual widgets that are either stateless (no internal state is changing) or stateful. They can be created from scratch to create unique looking visual components, or picked from a large set of widgets within the Material design (native android style) or the Cupertino design (native iOS style) [151]. In this way we can create native looking apps, however unlike Xamarin, whatever widget you build in the Dart code base will be rendered the same in iOS or Android.

We chose to follow the Cupertino style mainly due to the bottom placed tab bar.

Tier	Title	Sentence	Repl
0	Highest achievement reached	Wow! Congratulations, you reached the highest level in all life areas! You are showing a life that is in perfect balance	
1	Activity category hunter	Amazing, you succeeded to reach level 3 in the entire [Repl] side of the pyramide	{left, right}
2	You mastered an activity category	Your focus on [Repl] have payed of! You reached the highest level	An AC
3	You gained a level	Your [Repl] life is improving rapidly, You climbed levels! Congratulations	An AC
4	An accomplished day yesterday	Yay! You registered and rated [Repl] activities	A count of done activities
5a[a]	Routine activities	In the Routine tab you can plan activities that repeats every week or day. A great way to schedule a nice long bath every sunday	

[a] A total of 10 different quotes or hints exist in tier 5. One is selected by random.

Table 5.2: An overview of the implemented tier system.

As the app looks identical on both iOS and Android, we deliberately ignored the iOS or Android styled widgets but chose the most simple and usable design.

The implemented `MUBS` system was published to the Google Play store[11] and iOS app[12] store on February 2019.

5.3 Summary

Based on the UCD insights and the iterative design meetings, described in Chapter 4, we implemented two mHealth systems for BA. The two system were developed one and a half year apart, which resulted in separate implementation processes. A summary of the two systems is given in Table 5.3

[11] `https://play.google.com/store/apps/details?id=com.cachet.mubs01`
[12] `https://apps.apple.com/us/app/mubs-cachet/id1452891411`

	Application	
Setup	`Moribus`	`MUBS`
Google Play	✓[a]	✓
App store	✓	✓
Developing toolkit	Xamarin	Flutter
Programming environment	C#	Dart
Sensing platform	Sensus	CARP
Backend	AWS	Firebase
Hosted functions		GCP
Notification type	Local Push	Remote Push
Backend connection	Receive	Send & Receive

[a] `Moribus` was removed from Google Play due to the new regulations of data protection

Table 5.3: An overview of the `MUBS` and `Moribus` systems.

CHAPTER 6

Clinical Feasibility

6.1 Study Design

When we have to evaluate novel technology in its early stages it is argued that:

> "*A deep understanding of the how and why of the system use by its target users should be the central goal for evaluation of systems for health behavioral change*" [100, pp. 3069]

Therefore, following best practice within mHealth research in HCI, we ran two single-arm feasibility studies to evaluate `Moribus` and `MUBS`. For both systems we collected subjective and contextual data to investigate the usability, usefulness, and how it supports BA. Specifically, for `Moribus`, we wanted to understand the relation between activities and in what context they are done. These insights are valuable for therapists to trace how activities effect mood and management of the depressive symptoms.

`MUBS` was designed as a supportive system working independently from the clinic. As such, we were interested in understand how `MUBS` was used towards a better well-being. Furthermore, as the first example of a BA system providing personalized content-based recommendations, it would be interesting to know; (i) how the recommendations were *perceived*, (ii) whether they were *adopted*, and (iii) understand the specific *characteristic* behind the recommended activities that the patients chose to enact on.

The research protocol for the two studies were reviewed and approved by the Committee on Health Research Ethics of the Capital Region of Denmark (*j. 17018289*, and *H-19002943* respectively).

6.1.1 Participants

Usually, when conducting hypothesis-driven studies, the number of participants are determent through a power analysis. A analysis that estimates the number of samples needed to observe a statistical significant difference[1] that is deemed meaningful [106]. However, in a feasibility study we want to understand how the system functions in real life, so the recommended is to choose a number of participants similar to other

[1] also known as the effect size

published studies [117]. Guided by similar studies, we attempted to recruit within the range of 7-36 participants [20, 148, 157, 193, 15, 69].

Details on our recruitment process can be found elsewhere [B.1, B.2]. Briefly, clinical staff in Psychiatric Center Copenhagen at Rigshospitalet and the psychotherapeutic clinic at Nannasgade, Copenhagen N, informed eligible patients about the study. The information was written as a Q&A document, and available both as a flyer and online[2,3].

In `MUBS` we expanded the number of recruitment channels to include an advertisement on a national patient recruitment website[4], and a article on medium (Appendix D.2). Figure 6.1 schematically illustrates the number of patients recruited and continued for the duration of the studies.

We used the criterion's, as presented in Table 6.1, to assess the eligibility of the participants.

6.1.2 Apparatus

For the two studies we had to:

[2] `http://www.cachet.dk/research/studies/moribus`
[3] `http://www.cachet.dk/research/studies/mubs`
[4] `http://www.forsoegsperson.dk/show_ad.php?showit=2654&add=posted`

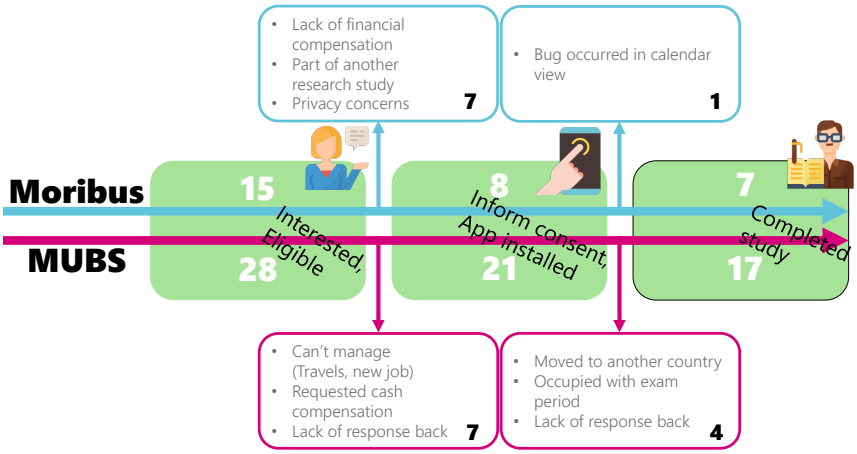


Figure 6.1: A visualization of the number of patients recruited and dropouts along the way.

Inclusion	Exclusion
Clinical diagnosed with an affective disorder	Age below 18
Experienced an depressive episode within the last year	Hospitalized
Own a Android or iOS smartphone	
Sufficient Danish or English language skills	
Previously or current enrollment in BA therapy[a]	

[a] Only applicable for the `Moribus` study

Table 6.1: An overview of the study criteria.

- Design and develop a mobile app system, which made use of BA techniques, backend services to store data, and high-order functions to derive statistical output and classifications.
- Program the systems within the confinement of the regulations defined by both Google Play and the iOS app store. Thereby, hosting the systems in a easy accessible public app repository for an easy installation process for the patients.
- Identify, obtain and implement questionnaires to support the field study. In our case, PHQ and WHO-5 to document any clinical changes during the study, as well as the Post-Study System Usability Questionnaire (PSSUQ) [108] and CACHET Unified Method for Assessment of Clinical Feasibility (CUMACF) [12] to measure the satisfaction with the usefulness and usability of the systems.
- Develop a set of unbiased informative questions for the end interview to assess the feasibility of the systems.

The studies were run on the patients own smartphone. The clinical change questionnaires were accessed through the app. For `Moribus` the questionnaires were accessed through a link to an online Google Forms survey, while in `MUBS` they were integrated in the app as a UI component. PSSUQ, and CUMACF were developed in Google Forms and send as an e-mail invitation to fill out the survey after the study. The end-interview was conducted in the clinic with a Nexus 5 Android phone running the voice recording app Otter[5]. In few cases the patients were unable to meet in the clinic. Here, we either called them directly through a mobile cell phone running Cube ACR[6] for voice recording, or Skype that has a built-in recording feature.

6.1.3 Procedure

We explained the study procedure for the systems in [B.1, B.2]. The main difference was due to `MUBS`'s design as a tool independent of the clinic. As a consequence, we

[5] `https://otter.ai/`
[6] `https://play.google.com/store/apps/details?id=com.catalinagroup.callrecorder`

did not invite the patients to the clinic for initial setup but added in-app tutorials, descriptions and help. An outline of the procedure is visualized in Figure 6.2.

6.2 Results

A summary of the demographics, usage data, and clinical findings for the two studies is provided in Table 6.2. We direct the reader to [B.1, B.2] for a thorough exposition of the results. Instead, we summarize the key findings for each study below.

6.2.1 Moribus

From the questionnaires, patients reported that they found `Moribus` useful to plan and register activities as part of BA therapy. When processing usage data, we found highly personalized patterns. In Figure 6.3 we visualize how the type of activities registered (Figure 6.3a) as well as the corresponding 'Pleasure' and 'Mastery' was different between patients (Figure 6.3b). Furthermore, both spatial and temporal differences existed in how patients chose to register activities. P82 was home on all occasions, while P88 registered 42% of the activities at home. Regarding time-of-day, same individualized behavior is seen as shown in Figure 6.4. Some chose a routine behavior of registering around the same time every day (P13, P92), while a more sporadic behavior was the case for others (P82, P88).

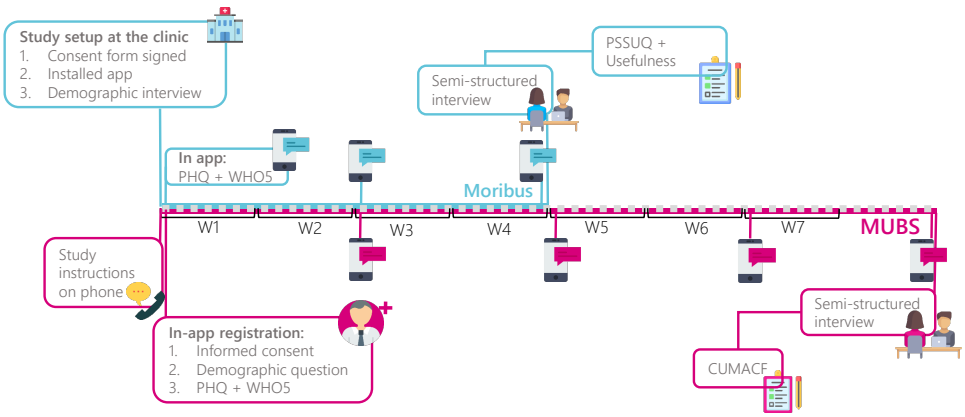
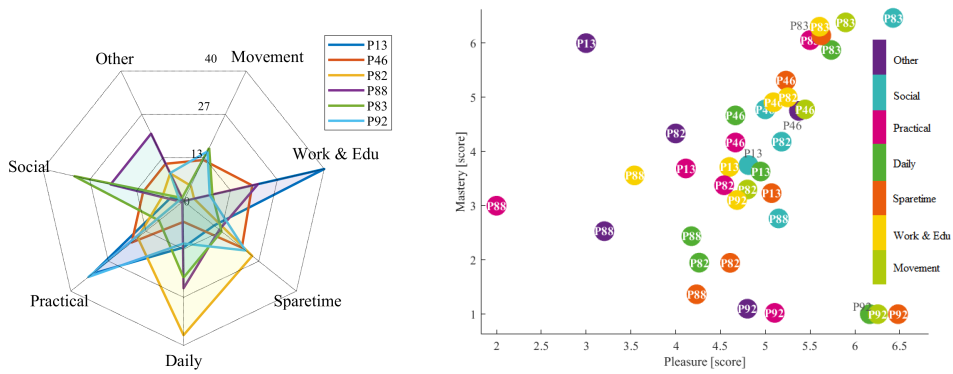


Figure 6.2: An overview of the study procedure. Every grey box represents one day. CUMACF: CACHET Unified Method for Assessment of Clinical Feasibility.

System	*Age*	*% Female*	*Compliance (%)*	*Total activities*	*From catalog (%)*	*Usability score*	*Init. PHQ*	*PHQ changed /week*
`Moribus`	43±20	57	71±16	1,684	25±34	3.25±1.15	5.57±4.35	1.67, p = .15
`MUBS`	33±10	82	76±15	2,895	62±30	3.83±1.08	10.06± 6.42	-1.09, p < .01

Table 6.2: A summary of demographics, usage data, and clinical outcome in the two studies. The calculated PHQ change is based on a repeated measures t test for the `Moribus` data (before and after). While `MUBS` used a Linear Mixed-Effects (LME) model to model all time points throughout the study.



(a) (b)

Figure 6.3: (a) The distribution of the different activity categories registered. The illustration has been copied from [B.1] (b) The average rating in 'Pleasure' and 'Mastery' for each activity category and each patient.

Through the semi-structured interviews we were particularly interested to know; (i) how did they find `Moribus` usable in comparison to their previous experience with BA paper-based tools; (ii) how they used the visual analytic tools and whether it provided any useful insights towards better self-efficacy.

The patients were positive of the transformation from paper to a digital system. They highlighted utilities such as the notifications, and the pie chart (Figure 5.1 A1) that made the experience better. P88 explained further how `Moribus` facilitated

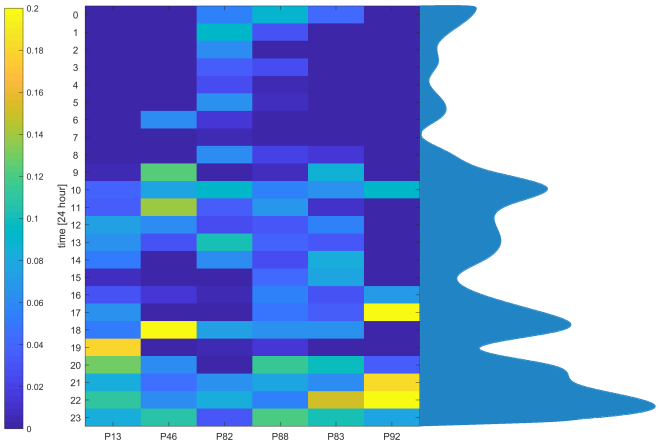


Figure 6.4: The distribution on when (24h) a patient went into `Moribus` and rated an activity as done, with a corresponding mood and 'Pleasure' score. The blue graph to the right illustrates the distribution across all patients.

registration in public:

> "*I don't want to bring pen and paper with me to register activities... and in front of people... on the phone it is much more anonymous*" [P88]

All patients agreed that the visual analytic tool was a vital asset over the paper-based version. As, a patient expressed:

> "*Under the statistics... I found the most valuable tool to analyze my activities. It provides an understanding of which activities helps me, and which gives problems that I need to be aware of - or completely avoid.*" [P82]

Of particular importance they highlighted the bar-chart illustrating the average 'Pleasure' and 'Mastery' for each activity category. They could then click to see what activities - within a category - that had an influence on the average. P82 explained how s/he became aware of the simple activity such as "*Drinking a cup of coffee*" had a positive impact. Although they were fond of the visual analytic tool, they concluded that the hourly activity registrations were tedious. The patients expressed it as more information flowing into the system than the other way around. However, as `Moribus` was designed to be part of BA therapy, it was an understandable reaction. In therapy sessions, based on the activity data they wrote down, they receive more direct support to plan specific activities. In the following study with `MUBS` we had the opportunity to explore whether this view on `Moribus` was reduced with a system

including a comprehensive list of inspirational activities, and a recommender model to provide actionable feedback.

6.2.2 MUBS

The system was perceived as highly usable. We can not directly compare `Moribus` and `MUBS`, since we did not run a Latin square design[7], but comments on usability was more pronounce in `MUBS`. It was repeatedly mentioned as an easy, straight-forward and visually aesthetic app with the simplified overview of planned and done activities:

> "*Layout-wise it is pleasant. You have the morning, afternoon, evening... Then you just simply add activities. I think it is nice and easily accessible, and it gives an overview that is not to confusion. I am insanely unorganized, and I need organization, so I really like to be able to access this page*" [P17]

We used a LME model to describe the time course of the submitted PHQ-8 score over depressive symptoms. The fitted model showed a statistical significant reduction in PHQ-8 that was weakened by a positive quadratic relation (Figure 6.5). The patients added on average 3.00 ± 1.65 activities to their planned overview, totalling 2,895

[7]The same patient testing both systems

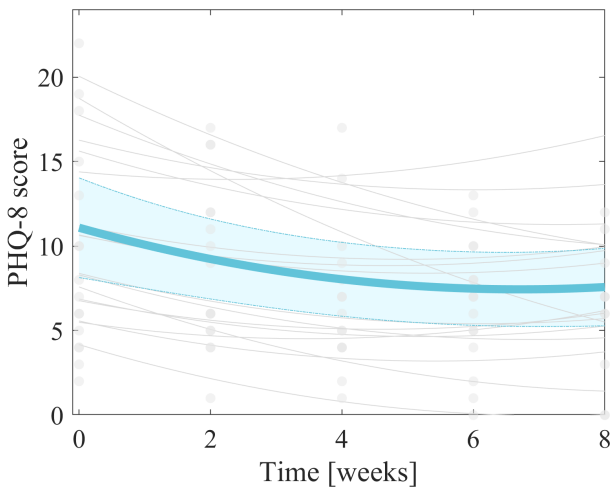


Figure 6.5: PHQ-8 as a function of time for the patients during the use of `MUBS`. The grey graph show each individual while the highlighted is the fitted model.

activities. A large majority was selected from the catalog of 384 pleasant activities (62%), while only 4% were picked from a personalized list of recommended activities. Of the activities picked from the recommendations, 83% were rated with a thumbs up, and a majority of these were difficulty 1 activities. The activities picked from the catalog or created by oneself, were in general of higher difficulty (Figure 6.6).

In BA therapy the clinician supports and educates the patient to plan activities. By having a schedule to follow it becomes easier for the patient to initiate the activity. Without the same type of support, `MUBS` was still able to support this behavior as 54% of all completed activities were planned. On their own, patients became aware of the benefit of planning:

> "*It worked best [`MUBS`] when I wrote it down either in the evening or the following morning, what I had to do that day. I experience a more efficient and better utilization of the day when I make the plan in advance*" [P11]

Therapist also supports the patient to identify pleasant activities. Similar behavior was supported by `MUBS` with the included catalog of pleasant activities, together with the required reflection by a thumbs- up or down. As a patient explained:

> "*When I have a bad day, I scrolled the list [catalog of activities] I realized that I had done good things such as applied creme on my hands, used lavender-oil or something like that. It was nice to see, and it contributed to the awareness that also such things are part of everyday life and take up time*" [P13]

In `Moribus`, the patient expressed the feeling that the system demanded more than it provided. No patient considered `MUBS` as a burden or similar in their everyday life. They emphasized the notification, and recommendations as source of feedback and support:

> "*I feel that the recommendations are relevant Its also the reason why I feel as if the app knows me. For instance, I have been very social lately, and suddenly I see that it recommends me to spend some time alone. I was surprised since I have been so active. In comparison to the first 14 days where I was alone, I was like, yes, that is totally fine, it is good to be alone sometimes. So that I have been*".

6.3 Summary

The clinical outcome of the two feasibility studies were as expected. `Moribus` was designed as a tool in BA therapy. It was given to patient to write and log activities on an hourly basis. As a tool, it did not have a clinical effect on depressive symptoms, but it functioned as intended. It had a high- usability and compliance rate during

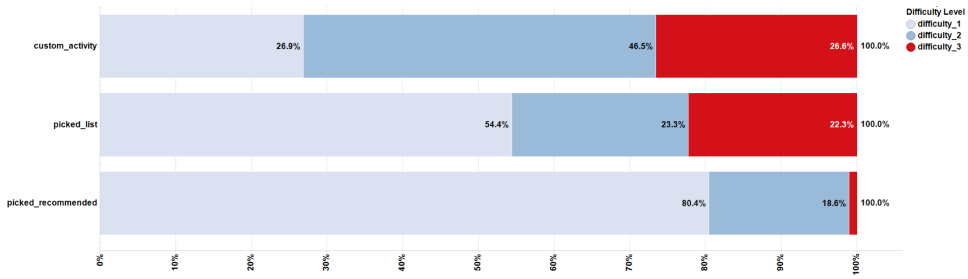


Figure 6.6: The difficulty picked when registering activities from either the recommendations, the list or custom.

the study, and the patient reported the activity registration as more convenient than the paper-based version. Since the activity was registered digitally, we added a visual analytic tool to show summary statistics of the registered activities. This functionality was positively received, providing value to the patients in the understanding of their behavior.

`MUBS` was designed as a stand-alone system to improve on well-being, without the involvement of the clinical personnel. The system had a significant clinical effect on depressive symptoms. On average the patient went from moderate levels of depression, to mild. They used the system as intended with more than half of all activities being planned beforehand. The catalog of 384 pleasant activities was highly praised and was a source of inspiration to schedule more enjoyable activities in the patients everyday life. The recommender model was also used as a source of inspiration, but less likely that the patient directly scheduled an recommended activity. The recommended activities were more applicable in situations where the patient had a larger free time-gap as a result of either unemployment or spare time from vacation or between studies.

CHAPTER 7

Discussion

This dissertation was motivated by the increased number of people experiencing depressive symptoms [149]. The increase in number of qualified clinical personnel does not follow same levels [26]. While we can not "copy" people, we can copy and scale-out digital solutions. Delivered care through smartphones has the potential to not only scale indefinitely, but also to reach suburbs without access to clinics, and provide free of charge help [133]. We saw a great potential to incorporate BA therapy into a mobile platform. BA is suitable since it requires minimal clinical contact, is straightforward to understand, and has proven as effective as other larger therapy methods [44]. Thus, to summarize our main research question, we were interested to answer:

> What is the feasibility of mobile and pervasive health technology for the treatment and care of depressive symptoms using a behavioral activation approach

We defined three sub-questions to guide our research work towards answering the main research question. We present each sub-question separate and discuss our findings and implications for future work.

> **RQ 1:** What is the relationship between mHealth sensor data and depressive symptoms? (Chapter 3)

In agreement with Torous et al. [187] we demonstrated the large potential that data collected from smartphones have to empower patients and clinicians with novel clinical tools. We saw that behavioral data carried strong personalized traits. For example, in the hand-written activity data, P5 had registered 55% of all activities as social-based, while P3 had 11% [A.1]. In the NB recommendation, the personalized models performed best with individual insights such as P1 that had lower preference for practical types of activities registered with words such as *housework, wash, clean, groceries, and cook* (Figure 3.5) [A.3]. Similar observation was seen in Springer et al. [177]. An expansion from a generalized model on behavior to individual regression models on a per user basis improved the explained variance from $R^2 = .434$ to $R^2 = .515$.

Set aside our data driven findings, personalization is already done in the clinic by the clinicians. It is demonstrated in various ways from fine-tuning treatment,

and inter-personal bonds, and selectively chose what information is more important than others [201]. Therefore, our results serve as a confirmation rather than being a unique finding – that personalization is also important when working with digital BA support.

Although individual differences exist in the relationship between behavioral data and depressive symptoms, there are still several connections that are applicable across patients. The therapist would never hesitate to recommend: "*Take a nice walk in your nearby park*", as getting fresh air, moving and interacting with other people are good things [122]. Mobile sensor data has the possibility to sense behavior in these domains. We identified seven behavioral domains (e.g., social, physical activity, and device interaction - Table 3.1) [A.2]. We found sensor data that showed consistent relationship with depressive symptoms across different studies:

- There was a positive relationship between depressive symptoms and the duration of smartphone usage
- The more the patient was staying home, as measured by the GPS sensor, the more severe depressive symptoms were reported
- Less physical activation of a patient, as measured by the accelorometer, resulted in higher levels of reported depressive symptoms

The information is easily captured by the smartphones sensors, and can be a great additional information for any therapeutic intervention on the phone. Particularly for BA, as these sensor measures directly link to behavior, and can be used to guide the therapy.

When modelling behavior one has to consider that there is numerous 'hidden' variables and relations. Different 'Pleasure' ratings for identical activities was seen in the activity data. In the statistical analysis of the hand-written data we introduced one type of second-order relation to the activity ratings. We showed that a day of routine activities affects the activity ratings positively for the next day [A.1, Figure 5]. Similarly, sleep [A.2] has an influence on mood and the perceived pleasure of doing things for several days, which is why self-assessments systems such as `Monarca` include tracking of these parameters [15]. Simple graphical representations on the combination of mood and sleep can then serve as support for behavioral sensemaking.

Understanding the type of data that relate to depressive symptoms suggest what behavioral data we should record and how it can guide therapy. The next step is to design the therapeutic intervention, as dictated by the next research question:

> **RQ 2:** What is the technical and UX design of mHealth for Behavioral Activation treatment of depressive symptoms? (Chapter 4-5)

Patients had several insightful opinions during the workshops and prototype sessions. In line with Saunders et al. [166], patients acknowledged that smartphone data matched with appropriate statistical methods can enhance clinical care and make

time spent with psychiatrists more efficient. The idea of supporting part of or complete therapeutic interventions on a digital system was positively received. Patients immediately recognized several advantages such as registrations in public, and the possibility of data-driven insights. Patients had several ideas on how to present the data to leverage these insights. It was emphasized that:

- An mHealth system should show the relation between activities, mood and/or 'Pleasure' rating

Additional abstract ideas were highlighted where mood is associated with the amount of different activities and how it is distributed. Visual analytic tools and sensor data to facilitate sensemaking has been found beneficial in many ways [98, 127]. It guides the patient to better understand his/her disorder, and provides a higher level of details in the recollection of past events which, in turn, improves satisfaction during therapy sessions.

Registering activities on a digital system was not considered a burden. The patients were willing to fill in activity information, even hourly-based as done on the standard paper based method [122], as long as it was done with few taps:

- Activity registration should be done with few taps

This led to the design choice of organizing activities into categories. Selecting an AC is a quick way of registering or planning activities instead of writing an activity text. The finalized AC design to label activities has been a recurrent motif in our work [A.1, A.3, B.1, B.2, C.1, C.2]. The change in terminology of American Time Use Survey (ATUS) was done to frame it in clinical terms that are understandable by clinicians. For instance, what ATUS defines as 'Personal care activities' with activities such as '*Sleeping*' and '*Eating & drinking*' is in mental health referred to as 'Daily living'. This category is particularly interesting within mental health research as it is the main factor affecting the quality of life [5]. Another example is the grouping of 'Leisure and sports' in ATUS. In our work to identify feature categories of objective data associated with depressive symptoms [A.2], physical activity was a highly studied feature (as we show in Table 3.1). It was consistently shown to be reduced in depression – also by another systematic review [169]. These considerations resulted in six AC's that carries clinical value. In our design we added an 'other' category to accommodate peculiarly activities as well as private behavior that the patient can choose to not impact the statistics from the six categories. We are aware that an 'other' categorization may lead to patients classifying activities into this category if in doubt. This can be avoided by designing custom categories as we saw in the EmotiCal system where the users were able to choose two custom labels [177]. However, we deliberately chose to keep a standard set of choices to be able to compare between participants as we did in [A.1] and to open op for the possibility of introducing novelty into a recommender feature as we demonstrated in [A.3]. In this case, unknown activities that might fit a given user, has to have predefined AC's within our design space when developing a content-based recommender system.

A patient mentioned the opportunity to:

- Receive inspirations of specific activities during planning

An opportunity that was welcomed by the clinicians as they tend to repeat similar activity suggestions for patients. As mentioned in Table 4.1 we included Lewinsohn & MacPhillamy [119] published list of pleasant activities as a result of their Pleasant Event Schedule (PES) research in 1982. Here, the participants were told to create a list of enjoyable activities, rank them, and introduce some of these activities in the following week plan. The list was published without any ranking, and we manually ranked all activities with a difficulty level. We discussed this choice with the patients, and they agreed that a standard difficulty level was fine as long as it can be edited. In our later empirical work, we learned that only few patients edited the level and were satisfied with the standard values [B.2]. The list also served an important purpose regarding our standardized AC. What is referred to as 'Recognition rather than recall', and is one of 10 usability heuristics [140]. The workload of the patient would be considerably lower by adding examples of activities within each AC. Although 'Social' activity might be easy to understand, a category such as 'Daily living' would be easy when the patient reads the suggested activities (e.g., "*get a warm shower*", or "*prepare a good meal*"). Adopting a list of specific activities is not uncommon in BA technology. Ly et al. [114] included a database of 54 activities that the patients could select from. They prefer to select from the list, but commented the activities as being too few and too simple [113]. From our empirical work we learned that a list of 384 activities was well received, and the simple activities were quickly filtered out by the difficulty level, when not applicable [B.2].

`MUBS` was designed as a standalone application to utilize BA methodology to assist and guide the user towards better well-being. As such, the workshop and the interdisciplinary design sessions, were focused on motivation, goal setting, feedback design. There was a shared theme centered around:

- positive reward for doing activities

One group in the workshop mentioned encouraging notifications such as: "*Yesterday you were really active*", "*Yesterday you took your medicine*", while another designed a trophy system. Interestingly, one patient shared an approach s/he used. Prior each day, the patient would spend the evening planning activities and assigning a number of stars to each. The goal was to collect as many stars as possible. Hence, more stars were allocated to tasks that were perceived as harder to initiate, as an incentive to perform them. The star system draws several parallels to the 'Mastery' score in BA [104]. The collection of points was adopted in our work [B.2] as the term 'Difficulty' points. Similar gamification alike features have been implemented in other well-being systems. For instance, Lin [109] used passive phone sensor to infer whether the participants behaviour was within recommended standards (e.g., if they slept 8 hours) and gave points accordingly. The point system was a motivating factor for the participant and enhanced the adherence to the system.

Goal setting is a concept within BA and an important technique for behavioral change [179]. Several standalone apps implements goal settings in various form. The PRIME app [168] had a list of 36 goals to chose from, e.g., *"deepen my relationship with my family"*, while in the HeadGear [46] app they had to reach the end of a 'road' through 30 daily tasks. Therefore there was an agreement that `MUBS` should include a goal in its design:

- Implement goal setting

The clinicians argued, as an app defined to improve well-being and reduce depressive symptoms, already had a self-appointed goal – *to improve well-being.* They referred back to the five general recommendations and said that balancing these should be a goal. The goal of a balance among the different areas led us to the design of a stair-case model for all AC. We refer to this star-case model as a pyramid [B.2].

The designed systems were then subject to a clinical feasibility study, leading to the third research question:

> **RQ 3:** What is the clinical feasibility of mHealth tools for Behavioral Activation treatment of depressive symptoms? (Chapter 6)

mHealth Technology as a tool for delivering or supporting therapy was a novel concept for many of the patients. Early interaction with a system is particularly difficult for patients experiencing depressive symptoms [159], which often results in poor adherence to mHealth systems [188]. Furthermore, the visual interface has to be tailored for a general audience with potentially limited numeracy and medical knowledge [76], and be engaging, interactive, and allow for personalization [79]. The number of factors, and barriers for the adoption of mHealth systems in everyday life are plenty [196, 192].

In our two studies we could relate to these barriers. In `Moribus` we designed hourly notifications to remind them and ease the process of registered 'done' activities. When inspecting the registration patterns (Figure 6.4), some users deliberately chose to register all their activities for the day in the evening. For them the hourly notifications were pointless. Minimizing all sorts of disruptive experiences is key for patient engagement [31]. This also includes bugs in the software. One patient stopped using `Moribus` after experiences errors in the calendar view. Some activity entries were reported as moved from one day to the next.

Allowing for personalization had the benefit of breaking down some of the previous listed barriers. The bubble chart in `Moribus` (Figure 5.1D3) displays the average 'Pleasure' for each AC as a function of time. The larger 'Pleasure', the larger bubble. One patient mentioned how it helped to understand that going to the fitness was most pleasurable in the evening, while another patient had no clue how to interpret the figure. Instead, the patient found the bar-chart (Figure 5.1 D1) insightful:

- The patients found their own personal way to enhance the benefit of the system

In the MUBS system, we saw how P5 used the daily reflection text-box that appears during mood rating. The patient was very happy to use this reflection as a diary that s/he could go back and see these small messages. A majority of the other patients never used this feature, but most importantly, they were not bothered by it. The same comment was given about other features in MUBS, including the Routine tab. Some patients mentioned how they currently did not have a daily life characterized by routine behavior but could see themselves use it once they had a job. Similarly they mentioned how the 'difficulty points' and the structure of the AC enhanced their use of MUBS to fit their personal situation and preferences:

- We experienced how the system allowed for personalization to improve on the usability

In MUBS, we documented a statistical significant clinical effect on depressive symptoms. The patients got better. While we cannot pinpoint the reason, or prove a causal effect, we can highlight some indicative data and subjective opinions.

The access to the 384 long catalog of pleasant activities was used by all patients in the study. They became aware of the smaller everyday activities that they actually engaged in, which is usually hidden in negative thinking [141]. The catalog also helped the patient to cope with periods lacking activities with positive enforcement:

- The catalog of pleasant activities facilitated planning of more enjoyable activities in the patients everyday life

In depressive periods, patient looses decision making abilities and ideas for enjoyable activities [26]. Several mental health apps lacks the ability to provide specific activities or actions to perform [109, 45, 46]. This is a significant weakness as specificity increases the probability of users enacting those activities [173]. We believed that the recommender model would be adopted more by the patients to schedule specific personalized activity. However, we experienced how the catalog was supporting this behavior, and the recommender model 'only', additionally to the catalog, added a sense of connection to the patient that the app was tailored to them. First, the recommendations were not applicable for patients that were experiencing less depressive symptoms and in general had a daily life characterized by routines. Second, when the patients were experiencing depressive symptoms, it was easier to passively scroll through the catalog than access the recommendations. Several patient expressed how they would have liked a notification that recommended a specific activity to schedule. This has also been proposed elsewhere [67]. While on the other side, in the workshops, they expressed less desire for an app to start telling them what to do. An idea, which we also touched upon briefly, would be to use the passive sensor data to 'trigger' a recommendation through notification, when the data suggest that the patient is experiencing more severe depressive symptoms.

Another facilitating aspect of the positive outcome from the app was the 'Difficulty points'. The patients explained how they edited the activity and increased the points to motivate the enacting of the activity. Once completed they felt rewarded,

and whatever behavior that followed was unimportant as their feeling of guilt was diminished.

- Collecting 'Points' and seeing the activity 'done', provided a feeling of reward and improved self-efficacy

Last, the mere fact of planning activities the night before, or in the morning, was repeatedly mentioned as most effective in regards to the utilization of the day. Without planning an activity, it was easier to postpone activities they had in mind. However, once it was written in the app they felt more obliged to follow the plan.

- By having a pre-scheduled plan of the day, was an incentive to follow it and carry out the activities

This discussion highlights the feasibility of `Moribus` and `MUBS` as systems for BA treatment of depressive symptoms. Although `MUBS` is designed to be used independent of the clinic, we want to make it clear that we do not – by any means – claim that `MUBS` can replace therapy sessions. Several of the patients had regular therapy session during the study, and several research studies repeatedly advocate for the combination of app usage and face-to-face therapy sessions [25]. However, `MUBS` might be a good system to fill the gap while patients are waiting for treatment [102]. Recently, it was shown that psychiatric patients that transfer from the kids & youth psychiatry clinic to the adult psychiatry unit are not sufficient supported and is often experienced as a no man's land [95]. Perhaps `MUBS` could be a prime candidate in this situation?

7.1 Limitations

We developed digital tools for patients with affective disorder, and collaborated with the Psychiatric Centre Copenhagen, Department of Affective Disorders, Rigshospitalet. Our inclusion criteria was patients who had experienced depressive symptoms. Surprisingly, a consequence of this was an overrepresentation of patients diagnosed with a bipolar disorder, both in the design meetings and the clinical evaluations. This led to more mood-related discussions during the end interview.

In `Moribus` we never evaluated the system with therapists. The system was designed as a tool to be used in a later consultation with a therapist, and – as a feasibility study – it would have been natural to investigate how they, together, used `Moribus`. Instead we kept the focus on the patient side of usage.

In `MUBS` we used time and day context-based filters. We did not include more informative context information, that was otherwise available from the passive sensor recordings, to post-filter the proposed activity recommendations. As an example of the consequence, a patient commented on a received recommended activity: "*go for a walk with the dog*", even though the patient did not have a dog.

Last, we would like to emphasize a limitation concerning potential privacy issues that the system has to be mindful of. In data-driven BA therapy with automatic

activity recommendations, there is a need to report daily life activities. Else, the system would have no idea what type of activities co-exist with your way of life, and the model will not be able to recommend appropriate activities. This type of intervention can open up to some privacy harms [175]. For this reason, there is a chance people might be concerned [162, 191]. The issue was out of scope for this dissertation but we welcome future work within ethical considerations in mHealth systems for BA.

7.2 Future work

This dissertation presents the design and development of a mHealth system adhering to the core BA ingredients. However there are still several treatment techniques in therapy [122] that are not implemented in our systems, which could add great value to our intervention. For instance, the therapists has techniques to break down patient barriers on initiating activities. They would assist the patient to break down the activity into smaller manageable sub-activities. To facilitate this support, we could add an option next to the delete button called "*Skipped*". The option could then initiate a dialog that helps the patient to define less difficult sub-activities and thereby reduce risk of hopelessness and discouragement.

As mentioned in the Limitations, the recommender model could be improved to include more context-based information. Although the model has to be aware that the intention of the patient might be to plan an activity for tomorrow, which represents a different context. Furthermore, several patients asked for a more accessible method to access or receive recommended activities. One method could be to directly attach the output of the recommender model (i.e., the probability of the patient enjoying an activity) on the catalog of activities. The patient can then sort and filter out the less relevant – as determent by the model – activities.

CHAPTER 8

Conclusion

The objective for this Ph.D. research project was to investigate the feasibility of a mHealth technology supporting Behavioral Activation (BA) methodology for the treatment of depressive symptoms. Through our theoretical contributions, we have shown how mHealth technology can be a perfect tool in depressive care and for BA data. It can assist to (i) utilize its embedded sensors to find generic correlations related to the patients' depressive symptoms, (ii) use computational power, to provide summary statistics of BA patterns, and (iii) use its connectivity capabilities to provide machine-learning based personalized recommendations calculated from cloud-based models.

We designed a mHealth smartphone app to support BA therapy, together with patients. The app was subject to a 4-week clinical feasibility study. We demonstrated that the patients were willing to use an app in their treatment, and that its visual analytic tools provided several insights into their behavior.

With the gathered activity data from the study we were able to develop and demonstrate a personalized recommender model. The model was able to suggest, with over 70% accuracy, specific activities that a patient would find enjoyable to do. We developed a second mHealth smartphone app that incorporates the personalized recommender model to provide BA treatment without the need of face-to-face therapy. Through a 8-week clinical feasibility study of the app, we saw high adherence to the principles of BA. The patients planned activities based on a list of general suggestions on pleasant activities, and from the recommended activities. The app improved their self-efficacy and awareness of positive behaviour that they enact on. They introduced more enjoyable activities in their everyday life and we observed a statistical significant reduction of the patients' depressive symptoms during the study.

A strength of this dissertation was our collaboration with the local psychiatric clinic. We had access to the same user group that we were building for. In this way, we were able to use clinical patients from the early design phase, to the final evaluation of the systems.

Based on our contributions, both theoretical and empirical, we can conclude that there is a strong incentive for mHealth technology in the treatment of mental disorders. Both as a tool to be used adjunct to therapy sessions, but definitely also as a private 'companion' that can support you ubiquitously.

APPENDIX A

Journal Papers

A.1 Data-driven learning in high-resolution activity sampling from patients with bipolar depression: Mixed-methods study.

JMIR MENTAL HEALTH Rohani et al.

Original Paper

Data-Driven Learning in High-Resolution Activity Sampling From Patients With Bipolar Depression: Mixed-Methods Study

Darius Adam Rohani[1,2], MSc(EE), Nanna Tuxen[3], Cand.Psych, Andrea Quemada Lopategui[1], MSc, Lars Vedel Kessing[3], DMSc, Jakob Eyvind Bardram[1,4], MSc, PhD

[1]Embedded Systems Engineering, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark
[2]Copenhagen Center for Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark
[3]Psychiatric Center Copenhagen, Rigshospitalet, Copenhagen, Denmark
[4]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Darius Adam Rohani, MSc(EE)
Embedded Systems Engineering
Department of Applied Mathematics and Computer Science
Technical University of Denmark
Richard Petersens Plads Building 324
Kgs. Lyngby, 2800
Denmark
Phone: +45 61412818 ext 45
Email: darohani@dtu.dk

Abstract

Background: Behavioral activation is a pen and paper-based therapy form for treating depression. The patient registers time activity locally, and together with the therapist, they agree on a plan to change behavior. Moreover, with the limited clinical contact, it is difficult to monitor patients' activity patterns and to detect relapse in depressive symptoms.

Objective: The objectives of this paper were to (1) exemplarily identify behavioral patterns through statistical analysis of the paper-based activity diaries, and (2) determine whether it is feasible to move the behavioral activation therapy format to a digital solution.

Method: We collected activity diaries from seven patients with bipolar depression, covering in total 2,493 hours of self-reported activity and 1,000 hours of sleep. Each hour of the day was reported for each activity. The activities were stipulated into 8 activity categories, and statistical analyses were conducted.

Results: Across all patients, movement-related activities were associated with the highest pleasure score followed by social activities. On an individual level, through a nonparametric Wilcoxon Signed-Rank test, one patient had a statistically significant larger amount of quiet-time activities when feeling bad (n=7, W=2.941, P=.003). Through a within-subject analysis of covariance, the patient tended to have a lower fun score on days he/she was depressed than manical/hypomanical days (n=7, F=3.231, P=.0.125). Furthermore, Granger causality tests indicated that two hours of daily social activity was optimal for the patients (F=7.556, 1 (SD=1.22), P=.25).

Conclusions: The data-driven statistical approach was able to find patterns within the behavioral charts that could assist the therapist as well as help design future technologies for behavioral activation.

JMIR Mental Health 2018;5(2):e10122; doi: 10.2196/10122

KEYWORDS

activities; behavior; behavioral activation; bipolar disorder; circadian rhythm; depression; hourly planning; psychotherapy; statistics

Introduction

Unipolar and bipolar depression has a high yearly prevalence in all age groups from children [1], adolescent [2] to elder adults [3,4]. The high prevalence imposes significant adverse consequences, including increased mortality, increased costs, loss of productivity, and lower well-being [5-8]. Treatment for bipolar depression consists of pharmacotherapy, psychotherapy,

https://mental.jmir.org/2018/2/e10122/

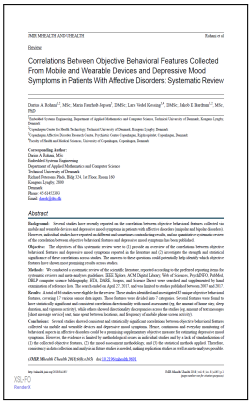
Authors

Darius Adam Rohani, Nanna Tuxen, Andrea Quemada Lopategui, Lars Vedel Kessing, & Jakob Eyvind Bardram

Journal

JMIR mental health, 5(2), e10122. (2018)

A.2 Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review



JMIR MHEALTH AND UHEALTH Rohani et al

Review

Correlations Between Objective Behavioral Features Collected From Mobile and Wearable Devices and Depressive Mood Symptoms in Patients With Affective Disorders: Systematic Review

Darius A. Rohani[1,2], MSc; Maria Faurholt-Jepsen[3], DMSc; Lars Vedel Kessing[3,4], DMSc; Jakob E Bardram[1,2], MSc, PhD

[1]Embedded Systems Engineering, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark
[2]Biomedicine Centre for Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark
[3]Copenhagen Affective Disorders Research Centre, Psychiatric Center Copenhagen, Rigshospitalet, Copenhagen, Denmark
[4]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Darius A Rohani, MSc
Embedded Systems Engineering
Department of Applied Mathematics and Computer Science
Technical University of Denmark
Richard Petersens Plads, Bldg. 324, 1st Floor, Room 140
Kongens Lyngby, 2800
Denmark
Phone: 45-51412385
Email: darohi@dtu.dk

Abstract

Background: Several studies have recently reported on the correlations between objective behavioral features collected via mobile and wearable devices and depressive mood symptoms in patients with affective disorders (unipolar and bipolar disorders). However, individual studies have reported indifferent and sometimes contradicting results, and no quantitative systematic review of the literature has been conducted to date to synthesize the reported correlations across the literature.

Objective: The objectives of this systematic review were to (1) provide an overview of the correlations between objective behavioral features and depressive mood symptoms reported in the literature and (2) investigate the strength and statistical significance of these correlations across studies. The answers to these questions could potentially help identify which objective features have shown most promising results across studies.

Methods: We conducted a systematic review of the scientific literature, reported according to the preferred reporting items for systematic reviews and meta-analyses guidelines. IEEE Xplore, ACM Digital Library, Web of Science, PsycINFO, PubMed, DBLP computer science bibliography, ETA, DiSE, Scopus, and Science Direct were searched and supplemented by hand examination of reference lists. The search was done on April 27, 2017, and was limited to studies published between 2007 and 2017.

Results: A total of 40 studies were eligible for the review. These studies identified and investigated 32 unique objective behavioral features, which were grouped into 6 categories: 3 on mobility, 2 on phone usage, 2 on sleep, 1 on vocal expression, and 1 on social. Of these, 8 were found to have statistically significant and consistent correlations (hereinafter, with mood) across studies, namely the amount of time spent in sleep duration, and vocal expressivity, while others showed directionality discrepancies across the studies (e.g. amount of text messages, [short average vs long] call, time spent between locations, and frequency of another phone screen activity).

Conclusions: Several studies showed consistent and statistically significant correlations between objective behavioral features collected via mobile and wearable devices and depressive mood symptoms. Hence, continuous and everyday monitoring of behavioral aspects in affective disorders could be a promising supplementary objective measure for estimating depressive mood. However, we identified a lack of standardization in the literature regarding (1) the definition of the behavioral features, (2) the collected objective features, (3) the mood assessment methodology, and (4) the statistical methods applied. Therefore, consensus on data collection and analysis in future studies is needed, making replication studies as well as meta-analyses possible.

(JMIR Mhealth Uhealth 2018;6(8):e165) doi:10.2196/mhealth.9681

http://mhealth.jmir.org/2018/8/e165/ JMIR Mhealth Uhealth 2018 - vol. 6 - iss. 8 - e165 - p.1
(page number not for citation purposes)

XSL•FO
RenderX

https://mhealth.jmir.org/2018/8/e165/

Authors
Darius Adam Rohani, Maria Faurholt-Jepsen, Lars Vedel Kessing, & Jakob Eyvind Bardram

Journal
JMIR mHealth and uHealth, 6(8), e165. (2018)

A.3 Recommending Activities for Mental Health and Well-being: Insights from Two User Studies

SPECIAL SECTION ON NEW FRONTIERS IN COMPUTING FOR NEXT-GENERATION HEALTHCARE SYSTEMS 1

Recommending Activities to Improve Mental Health and Well-being: Insights from Two User Studies

Darius A. Rohani, Member, IEEE, Aaron Springer, Jakob E. Bardram, and Steve Whittaker,

Abstract—Engaging in daily activities that engender positive affect (e.g., exercise and socializing) are critical for emotional well-being. However current mental health interventions have not exploited recommender approaches to encourage such behaviors. We, therefore, have developed two mobile applications that recommend activities to users. We report on two studies of these applications. The first developed mobile applications allowing us to collect high-quality data about explicit activities from two populations: a clinical sample diagnosed with a mood disorder (n = 21 with interviews) and a non-clinical sample (n = 18 with interviews). Activities were labeled with a type (e.g., social, leisure, work), and rated for their impact on mood. We used a personalized multivariate Bayesian linear baseline (MB) and a Support Vector Machine (SVM) to model the activities as a function of mood in a pooled mood outcome. We evaluated the analysis using 10-fold cross validation. We found that the personalized MB model significantly outperformed the SVM model. In the clinical sample, the comparison was a significant difference between the two models, $Z(N=18)=2.0$, $P(R)=0.02$ and $Z(N=18)=2.0$, $P(R)=0.02$ and the personalized model also outperformed the non-personalized model when collecting $50\,(SE=20)$ activities. This research sheds light on recommendations for mental health, showing that personalization is key to recommend the right activity to each user.

Index Terms—Recommendation; Activity; Well-being; behavior; context-based; m-health; machine learning

1 INTRODUCTION

Mental health and well-being are major public and personal health issues, being a leading cause of disability and disease burden worldwide [1]. 30% of men and 40% of women experience a major depressive episode at least once in their life with sometimes being even more severe for people with mental illnesses [2]. Depression is as a cause of disease burden [2], and many people remain untreated [4]. As a result, there is a growing research effort to develop personal informatics technologies to alleviate mental health symptoms (see [5], [6] for reviews). These systems have been developed for a variety of mental illnesses [7], [8], [9], [10], [11]. Nevertheless there are important limitations to these existing applications: they often rely on manual user tracking of symptoms and interventions; strategies are sometimes based around simple self-help heuristics. Together these limitations may explain relatively low levels of long-term adoption [12], [13]. The current research aims to address these limitations by exploring systems that were also driven by simple heuristics. Some, of course, many recommender platforms reduce the analytic burden for users by providing explicit product recommendations, for books, movies or songs. But not only increasing platform usability, but also increasing product sales [14].

This paper explores the application of recommender techniques to the mental health domain. The key insight underlying this work is that when patients regularly enact activities that they rate as pleasant, then has a significant positive impact on mental health [16]. A successful behavioral intervention is therefore to recommend personal activities that are likely to improve the patient's affective state. For example, walking the dog or meeting friends consistently rated as pleasant. In fact, a behavioral activation (BA) intervention has been demonstrated as a effective treatment for depression, comparable to larger psychotherapy methods such as Cognitive Behavioral Therapy (CBT) for alcohol and drug therapies [17], [18].

However, current mental PHI interventions rely on some digital approaches where participants manually log their daily behaviors for several weeks. Patients then collaborate with their therapist to analyze this record, to identify 'pleasant' events. Patient and therapist then generate a behavioral plan involving a weekly schedule of those pleasant events. However, there are drawbacks to this approach. First, it is time consuming, requiring several hours of work, may be prone to errors. Third, it also time and cost-involved in scheduling therapist meetings. As a result, patients are receptive to computational approaches with surveys showing 91% of patient respondents are positive about computer-based tools, and 80% believe they would be able to use them control their other [19]. An app can be accessed privately without needing to disclose personal details or schedule medical health services [20].

The goal of the current research is therefore to explore how to develop recommendations for BA using a mobile app. Specifically, we examine whether recommender methods might allow us to reliably predict mood-boosting events. Generating reliable predictions would allow us to use these models to recommend pleasant activities to users.

This study uses data collected from two mobile applica-

• D. A. Rohani and J. E. Bardram are with the Department of Health Technology, The Technical University of Denmark, Denmark.

• A. Springer and S. Whittaker are with the University of California Santa Cruz, USA.

Manuscript received July 12, 2019

Under review

Authors

Darius Adam Rohani, Aaron Springer, Victoria Hollis, Jakob Eyvind Bardram, & Steve Whittaker

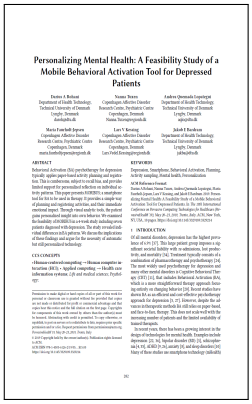
Journal

IEEE Transactions on Emerging Topics in Computing. Special Section on New Frontiers in Computing for Next-Generation Healthcare Systems, 8(1), (2020)

APPENDIX B

Conference Papers

B.1 Personalizing mental health: A feasibility study of a mobile behavioral activation tool for depressed patients



Personalizing Mental Health: A Feasibility Study of a Mobile Behavioral Activation Tool for Depressed Patients

Darius A. Rohani
Department of Health Technology
Technical University of Denmark
Lyngby, Denmark
darohani@dtu.dk

Nanna Tuxen
Copenhagen Affective Disorder Research Centre, Psychiatric Centre Copenhagen, Denmark
Nanna.Tuxen@regionh.dk

Andrea Quemada Lopategui
Department of Health Technology
Technical University of Denmark
Lyngby, Denmark
aqlop@dtu.dk

Maria Faurholt-Jepsen
Copenhagen Affective Disorder Research Centre, Psychiatric Centre Copenhagen, Denmark
maria.faurholt-jepsen@regionh.dk

Lars V. Kessing
Copenhagen Affective Disorder Research Centre, Psychiatric Centre Copenhagen, Denmark
Lars.Vedel.Kessing@regionh.dk

Jakob E. Bardram
Department of Health Technology
Technical University of Denmark
Lyngby, Denmark
jakba@dtu.dk

ABSTRACT

Behavioral Activation (BA) psychotherapy for depression typically applies paper-based activity planning and reporting. This is cumbersome, subject to recall bias, and provides limited support for personalized reflection on individual triggers. We have developed a mobile version of the paper-based tool that BA is to be used in therapy. It provides a simple way of planning and registering activities, and their immediate emotional impact. Through visual analytic tools, the patient gains personalized insight into own behavior. We examined the feasibility of the mobile tool in a study with 10 outpatients diagnosed with depression. The study revealed individual differences in BA patterns. We discuss the implications of these findings and argue for the necessity of automatic but still personalized technology.

CCS CONCEPTS

• Human-centered computing → Human computer interaction (HCI); • Applied computing → Health care information systems; Life and medical sciences; Psychology.

KEYWORDS

Depression, Smartphone, Behavioral Activation, Planning, Activity sampling, Mental health, Personalization

ACM Reference Format:

Darius A Rohani, Nanna Tuxen, Andrea Quemada Lopategui, Maria Faurholt-Jepsen, Lars V. Kessing, Jakob E. Bardram. 2019. Personalizing Mental Health: A Feasibility Study of a Mobile Behavioral Activation Tool for Depressed Patients. In: 19th International Conference on Pervasive Computing Technologies for Healthcare (Pervasive-Health'19), May 20–23, 2019, Trento, Italy. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3329214.3329214

https://dl.acm.org/citation.cfm?id=3329214

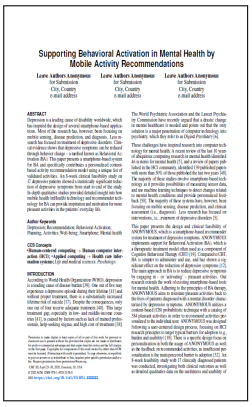
Authors

Darius Adam Rohani, Nanna Tuxen, Andrea Quemada Lopategui, Maria Faurholt-Jepsen, Lars V Kessing, & Jakob Eyvind Bardram.

Proceedings

Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare, Pervasive-Health'19, pp. 282-291. (2019)

B.2 Supporting Behavioral Activation in Mental Health by Mobile Activity Recommendations



Supporting Behavioral Activation in Mental Health by Mobile Activity Recommendations

Leave Authors Anonymous for Submission City, Country e-mail address

Leave Authors Anonymous for Submission City, Country e-mail address

Leave Authors Anonymous for Submission City, Country e-mail address

ABSTRACT

Depression is a leading cause of disability worldwide, which has inspired the design of several smartphone-based applications in the HCI community, ranging from mood tracking, mobile sensing, disease prediction, and diagnosis. Less research has focused on treatment of depressive disorders. Clinical evidence shows that depressive symptoms can be reduced through behavior change – a method known as Behavioral Activation (BA). This paper presents the design and clinical feasibility of BA and specifically contributes a personalized, computer-based activity recommendation model using a unique list of validated activities. An 8-week clinical feasibility study on 17 depression patients showed a statistically significant reduction in depressive symptoms and high usability of the system. In-depth qualitative studies provided detailed insights into how mobile health (mHealth) technology and recommender technology for BA can provide inspiration and motivation for more pleasant activities in the patients' everyday life.

Author Keywords

Depression; Recommendation; Behavioral Activation; Planning; Activities; Well-being; Smartphone; Mental health

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); •Applied computing → Health care information systems; Life and medical sciences; Psychology;

INTRODUCTION

According to World Health Organization (WHO), depression is a leading cause of disability worldwide [1]. About 10% experience a depressive episode during their lifetime [2] and without proper treatment, there is a substantially increased lifetime risk of suicide [37]. Despite the consequences, only one out of four receive adequate treatment [40]. This large treatment gap, which is even larger in low-income countries [41], is caused by factors such as lack of trained professionals, help-seeking stigma, and high cost of treatment [34].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI'20, April 25–30, 2020, Honolulu, HI, USA

© 2020 ACM 978-1-4503-6819-3/20/04 ...$15.00

DOI https://doi.org/10.1145/3313831.3376311

The World Psychiatric Association and the Lancet Psychiatry Commission have recently argued that a drastic change in mental healthcare is needed and points out that the only treatment recommendation for depression is medication, unlike psychiatry, which they refer to as a Digital Psychiatry [2].

These challenges have inspired research into computer technology for mental health. A recent review of the last 35 years of ubiquitous computing research in mental health identified six topics for mental health [17], and a review of papers published in the HCI community, identified 19 published papers on mobile mental health interventions in the last decade [6]. The majority of these studies involve smartphone-based technology as it provides possibilities of measuring sensor data, and use machine learning techniques to detect changes related to mental health conditions and provide personalized feedback. The HCI research has mostly focused on mobile sensing, focusing on mobile sensing, disease prediction, and clinical assessment (i.e., diagnosis). Less research has focused on interventions, i.e., *treatment* of depressive disorders [5].

This paper presents the design and clinical feasibility of ANONYMOUS, which is a smartphone-based recommender system for treatment of depressive symptoms. ANONYMOUS is based on Behavioral Activation (BA), which is a well-known therapeutic treatment model, often used as a component of Cognitive Behavioral Therapy (CBT) [19]. Comparatively CBT, BA is simpler to administer and use, and has shown a significant effect on the reduction of depressive symptoms [12]. BA is a method for CBT, which aims to increase well-being by engaging in – or 'activating' – pleasant activities. Our research extends the work of existing smartphone-based tools for mental health. Adhering to the principles of BA therapy, ANONYMOUS aims to minimize pleasant activities back to the patient's everyday life, which have been reported to be reduced in depressive symptoms. ANONYMOUS utilizes a content-based CTR probabilistic technique with a catalog of 544 pleasant activities in order to recommend activities personalized to the individual user. ANONYMOUS was designed and developed following a user-centered approach, with research principles to adapt typical barriers for adaptation (e.g., burden and usability) [18]. There is a specific design focus on personalization in both the usage of ANONYMOUS as well as the feedback on recommended activities, as ineffective personalization has been reported as a barrier for use [18]. An 8-week feasibility study with 17 clinically diagnosed patients was conducted, investigating both clinical outcomes as well as detailed qualitative data on the usefulness and usability of

Under review

Authors

Darius Adam Rohani, Andrea Quemada Lopategui, Nanna Tuxen, Alban Maxhuni, Maria Faurholt-Jepsen, Lars V Kessing, & Jakob Eyvind Bardram.

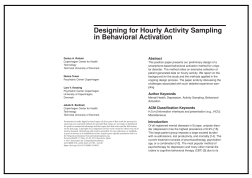
Proceedings

Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI'20, pp. 1-10. (2020)

APPENDIX C

Workshop Papers

C.1 Designing for hourly activity sampling in behavioral activation



Designing for Hourly Activity Sampling in Behavioral Activation

https://dl.acm.org/citation.cfm?id=3154919

Authors

Darius Adam Rohani, Nanna Tuxen, Lars Vedel Kessing, & Jakob Eyvind Bardram.

Proceedings

Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, Pervasive-Health'17, pp. 431-435. (2017)

C.2 Supporting smartphone-based behavioral activation: A simulation study

UbiCom'17 ISWC'17 SEPTEMBER 11–15, 2017 Maui, HAWAII, USA

Supporting Smartphone-based Behavioral Activation: A Simulation Study

Jakob E. Bardram,
Darius A. Rohani,
Center for Sensorimotor Interaction, Dept. of Health Technology, Technical University of Denmark, Ørsted Plads, Building 349, 2800 Lyngby, Denmark
{jakob,darohani}@dtu.dk

Nanna Tuxen,
Maria Faurholt-Jepsen,
Lars V. Kessing,
Copenhagen Affective Disorder Research Centre (CADIC), Psychiatric Center Copenhagen, Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark
{nanna.tuxen,maria.faurholt-jepsen,lars.vedel.kessing}@regionh.dk

ABSTRACT
Bipolar disorder affects over 60 million people worldwide. Among the depressive episodes, the individual plays an extensive collection of patterns repeated weekly, often on a 24-hour basis, after an episode has ended. Such patterns are important to monitor and understand in order to prevent future episodes. This paper presents a simulation study of a smartphone-based behavioral activation system, which aims at helping people with bipolar disorder to plan, monitor and reflect upon their daily activities, and thereby prevent future episodes. The study shows that the system can detect and predict patterns in activity data, and that it can prevent 12 episodes of the data and alleviate the repercussions for the remaining episodes. We discuss implications for behavioral interventions.

CCS CONCEPTS
• **Applied computing** → **Psychology and mental health**; **computing systems and business systems**; **Health informatics**; **Decision analysis**;

https://dl.acm.org/citation.cfm?id=3125617

Authors
Jakob Eyvind Bardram, Darius Adam Rohani, Nanna Tuxen, Maria Faurholt-Jepsen, & Lars V. Kessing.

Proceedings
Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, UbiComp '17 (2017)

APPENDIX D

Popular science

D.1 Din personlige digitale terapeut

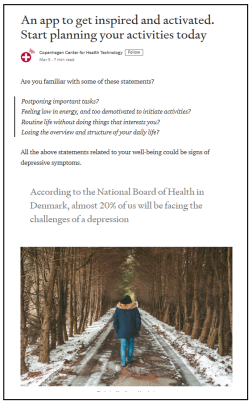
Din personlige digitale terapeut

https://ipaper.ipapercms.dk/TechMedia/Medicoteknik/2019/4/?page=10#/

Authors
Darius Adam Rohani

Journal
Medikoteknik, 4, pp. 10-11. (2019)

D.2 An app to get inspired and activated. Start planning your activities today



An app to get inspired and activated. Start planning your activities today

Copenhagen Center for Health Technology FOLLOW
Mar 6 · 7 min read

Are you familiar with some of these statements?

Postponing important tasks?
Feeling low in energy, and too demotivated to initiate activities?
Routine life without doing things that interests you?
Losing the overview and structure of your daily life?

All the above statements related to your well-being could be signs of depressive symptoms.

According to the National Board of Health in Denmark, almost 20% of us will be facing the challenges of a depression

https://medium.com/@cph.cachet/an-app-to-get-inspired-and-activated-start-planning-your-activities-today-8da835d97eac

Authors
Darius Adam Rohani

Journal
Medium.com, 7min read. (2019)

APPENDIX E

Workshop agenda example

Workshop [MUBS]
Part 1. Registrations within the app

DARIUS A. ROHANI, ANDREA QUEMADA LOPATEGUI, NANNA TUXEN, MARIA FAURHOLT-JEPSEN, LARS V. KESSING, JAKOB E. BARDRAM
Psychiatric Center Copenhagen, Department O, 6233, Rigshospitalet.
March 15, 2018

Setting

The workshop room will be decorated with inspirational images on the wall and images of activity categories. The images are from CACHET's stock database, and features people, nature, achievements etc (see Figure 1). There will be one facilitator (DAR) and one person documenting (AQL). The participants includes a software engineer, UX designers, a psychologist, a medical doctor, and three patients diagnosed with bipolar disorder.

Coffee and cookies/healthy-alternatives will be prepared (DAR).

Figure 1: Inspiration pictures

Music will be played during the workshop. A calming inspirational theme that boosts the creativity of the participants.

Goal

The aim of this workshop is to gather knowledge about how the users of the app (app name: `MUBS`) should register their daily tasks. The daily tasks are defined as five + one activities that the users should do every day:

1. Get a good sleep
2. Take your medicine
3. General Hygiene
4. Get outside
5. Get in contact with someone else (social)
6. Register your mood, in the end of the day

To achieve the task, the workshop will be facilitated with two main activities. First an active idea creation by the attendees, then a pros and cons on existing idea sketches developed by DAR, AQL, JEB. The following section gives a detailed description of the agenda.

Agenda

1. Cachet introduction

Author Jakob E. Bardram

Duration 3 minutes

Jakob will re-use his old introductory presentation he gave the last workshop. It introduces an outline of the day, the goal, existing work and inspirational slides of other health-related applications.

This is followed by each participant given a short presentation of their hobbies and interests. A name tag is created and applied.

2. Behavioral Activation & `MUBS`

Author Darius A. Rohani

Duration 5 minutes

Behavioral Activation (BA) is introduced. Then an introduction is given about our previous/on-going BA app `MORIBUS`, with a talk of the things that is challenging, what we learned, did not work, and needs to be changed. This is then ended with a presentation of `MUBS`, with the five + one activities.

3. Inspirational break

Author Darius A. Rohani

Duration 4 minutes

The participants will make a short break, stand up, and walk among the inspirational images that are placed on the walls around the room. Music will be playing. (DAR) will explain this: "now stand up, look at the pictures around the room, do you like it? Does it inspire you to be more active? Do you want to be in that situation?". "Now we sit down and start the first activity of the workshop".

4. Hands-on

Author Darius A. Rohani

Duration 2*15 minutes

Then we split in three groups, drawing a number from a bag, with the participants given several semi-empty smartphone sketches as illustrated in Figure 2

Visualization

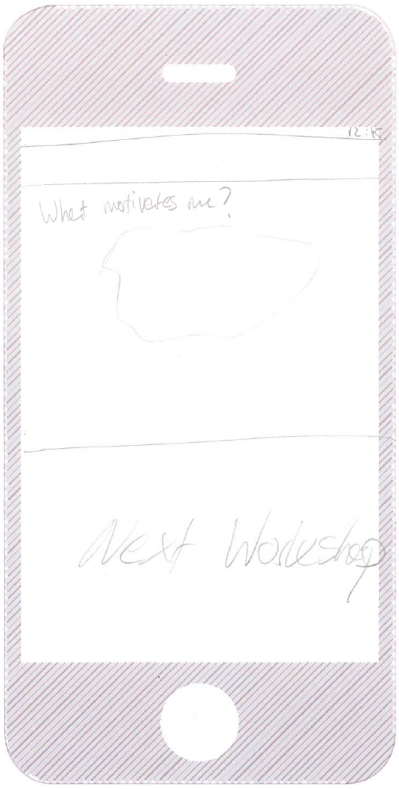
The groups are asked to sketch two ideas/ways of illustrating - on the main app screen - how they are progressing with the five + one activities. They should have in mind that the goal of the visualization is to motivate the user to do the activities.

They should think about two main things: The type of visualization (graph, butterfly, flowers) and the data shown (last 7 days of the registration of the activities - "*Did I go out the door on Tuesday?*", the amount of activities done to reach a goal - "*I did 3 out of the 6 things today, yeah!*", the registered time or context - "*I already went out the door at 12 yesterday, nice*").

Registration

The attendees are asked to sketch two ideas of how the users are suppose to interact with the main screen to register that they did do the five + one activities. The goal of this interaction should be an immediate feeling of

Figure 2: A semi-empty phone sketch



achievement: *Yes, I did it*. Of course both the Registration and Visualization parts and should be combined.

The ideas are shared and discussed among the participants, in order to get all the thoughts out and learn as much as possible from each other.

The 'hands-on' is ended.

5. 2nd break - energizer

Author Darius A. Rohani

Duration 4 minutes

Everyone close their eyes. We play some relaxing music (AQL). The facilitator describes a very relaxing scenario for 1 minute. The second minute the facilitator starts describing really extravagant objects in the relaxing

scene. The participants are not longer relaxed and the facilitator is boosting their creativity. The facilitator asks the attendees to open their eyes, stand up, raise up their hands and let them fall x 3 times. Attendees are energized and ready for the next activity.

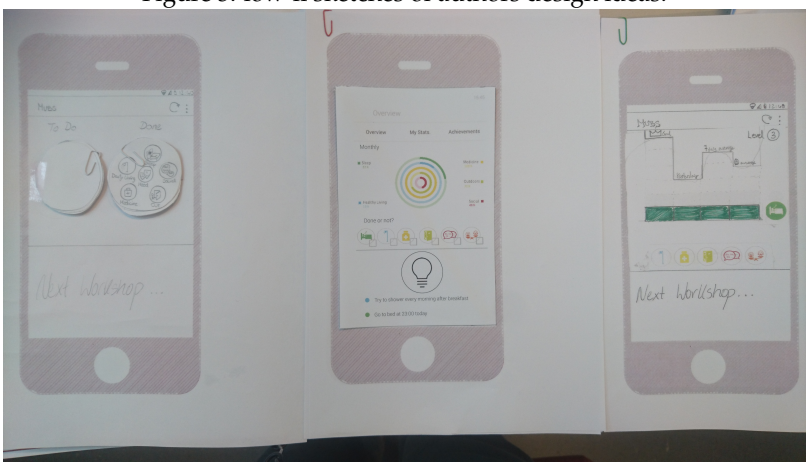
4. Attendee comments

Author Darius A. Rohani

Duration 15 minutes

Three new groups are formed by facilitator giving a number from 1-3 to each participant. They are presented with 3 sketches, as visualized in Figure 3, of ideas that have been developed by authors (DAR), (AQL), (JEB). One sketch for each group, which rotates 2 times. The sketch is placed on a A3 sheet, and each group is asked to write pros/cons and idea comments down for each sketch on the A3 sheet.

Figure 3: low-fi sketches of authors design ideas.



5. New sketchings [BONUS]

In case of more time, the attendees are handed out clean/empty phone sketches and are completely free to draft ideas down together with the participants.

APPENDIX F

Sensing toolkits

Table F.1 provides a quick overview of the available mobile sensing toolkit during the implementation of `Moribus`.

Name	Organization	Lang.	Last update	Platform	GPS	Phone log	Screen events
Funf	Massachusetts Institute of Technology (MIT)	Java	11 Jan 2016	Android	✓	✓	✓
Sensingkit	Queen Mary University of London	Java/Obj.C	12 Nov 2018	Android, iOS	✓	✓	✓
Sensus	University of Virginia	C#	Up-to-date	Android, iOS	✓	✓	✓
AWARE	University of Oulu	Java/Obj.C	Up-to-date	Android, iOS	✓	✓	✓
PACO	Google	Java/Obj.C	Up-to-date	Android, iOS			
EmotionSense	University of Cambridge	Java/Swift	20 Apr 2016	Android, iOS	✓	✓	✓
ohmage	University of California, Los Angeles	Java/Obj.C	6 Aug 2018	Android, iOS	✓		
Purple Robot	Northwestern University	Java	11 Apr 2016	Android	✓	✓	
Research Stack	Cornell University	Java	27 Jan 2017	Android			
Open data kit	University of Washington	Java	Up-to-date	Android			

Table F.1: A list of free public toolkits for mobile sensing.

Scientific Contribution

[A.1] **Darius A Rohani**, Nanna Tuxen, Andrea Quemada Lopategui, Lars Vedel Kessing, and Jakob Eyvind Bardram. Data-driven learning in high-resolution activity sampling from patients with bipolar depression: Mixed-methods study. *JMIR mental health*, 5(2):e10122, 2018.

[A.2] **Darius A Rohani**, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E Bardram. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review. *JMIR mHealth and uHealth*, 6(8):e165, 2018.

[A.3] [SUBMITTED] **Darius A Rohani**, Aaron Springer, Victoria Hollis, Jakob E Bardram, and Steve Whittaker. Recommending activities for mental health and well-being: Insights from two user studies. *IEEE Transactions on Emerging Topics in Computing. Special Section on New Frontiers in Computing for Next-Generation Healthcare Systems*, 8(1):1–12, 2020.

[B.1] **Darius A Rohani**, Nanna Tuxen, Andrea Quemada Lopategui, Maria Faurholt-Jepsen, Lars V Kessing, and Jakob E Bardram. Personalizing mental health: A feasibility study of a mobile behavioral activation tool for depressed patients. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth'19, pages 282–291, New York, NY, USA, 2019. ACM.

[B.2] [SUBMITTED] **Darius A Rohani**, Andrea Quemada Lopategui, Nanna Tuxen, Alban Maxhuni, Maria Faurholt-Jepsen, Lars V Kessing, and Jakob E Bardram. Supporting behavioral activation in mental health by mobile activity recommendations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 3:1–3:13, New York, NY, USA, 2019. ACM.

[C.1] **Darius A Rohani**, Nanna Tuxen, Lars V. Kessing, and Jakob E. Bardram. Designing for hourly activity sampling in behavioral activation. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth '17, pages 431–435, New York, NY, USA, 2017. ACM.

[C.2] Jakob E. Bardram, **Darius A Rohani**, Nanna Tuxen, Maria Faurholt-Jepsen, and Lars V. Kessing. Supporting smartphone-based behavioral activation: A simulation study. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, UbiComp '17, pages 830–843, New York, NY, USA, 2017. ACM.

Bibliography

[1] Saeed Abdullah, Mark Matthews, Ellen Frank, Gavin Doherty, Geri Gay, and Tanzeem Choudhury. Automatic detection of social rhythms in bipolar disorder. *Journal of the American Medical Informatics Association*, 23(3):538–543, 2016.

[2] Danish Energy Agency. Telestatistik - andet halvår 2018. Homepage, December 2018. Retrieved July 17, 2019 from `https://ens.dk/sites/ens.dk/files/Tele/telestatistik_-_andet_halvaar_2018.pdf`.

[3] Jorge Alvarez-lozano, Mads Frost, Venet Osmani, Jakob Bardram, Lars Vedel Kessing, Oscar Mayora, and Maria Faurholt-Jepsen. Tell me your apps and I will tell you your mood: correlation of apps usage with bipolar disorder state. *7th International Conference on Pervasive Technologies Related to Assistive Envotinments (PETRA)*, 1(212), 2014.

[4] American Pyschiatric Association (APA). Practice Guidelines for the Treatment of Patients With Major Depressive Disorder. *American Pyschiatric Association (APA)*, 150(4):1–26, 2010.

[5] Christian K Andersen, Kim U Wittrup-Jensen, Anette Lolk, Kjeld Andersen, and Per Kragh-Sørensen. Ability to perform activities of daily living is the main factor affecting quality of life in patients with dementia. *Health and quality of life outcomes*, 2(1):52, 2004.

[6] Emily Anthes. Mental health: there's an app for that. *Nature News*, 532(7597):20, 2016.

[7] Patricia A. Arean, Kevin A. Hallgren, Joshua T. Jordan, Adam Gazzaley, David C. Atkins, Patrick J. Heagerty, and Joaquin A. Anguera. The use and effectiveness of mobile apps for depression: Results from a fully remote clinical trial. *Journal of Medical Internet Research*, 18(12):1–13, 2016.

[8] Joost Asselbergs, Jeroen Ruwaard, Michal Ejdys, Niels Schrader, Marit Sijbrandij, and Heleen Riper. Mobile Phone-Based Unobtrusive Ecological Momentary Assessment of Day-to-Day Mood: An Explorative Study. *Journal of Medical Internet Research*, 18(3):e72, 2016.

[9] American Psychiatric Association et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[10] Per Bækgaard, John Paulin Hansen, Katsumi Minakata, and I Scott MacKenzie. A fitts' law study of pupil dilations in a head-mounted display. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, page 32. ACM, 2019.

[11] Jakob E Bardram. The art of doing a phd. Homepage, June 2016. Retrieved July 01, 2019 from `http://www.bardram.net/wp-content/uploads/2017/12/Art.PhD_.DTU_.v.30.pdf`.

[12] Jakob E. Bardram. CACHET Unified Methodology for Assessment of Clinical Feasibility. Technical report, Copenhagen Center for Health Technology, Copenhagen , Denmark, 2018. Available from http://www.cachet.dk/research/cumacf.

[13] Jakob E. Bardram. Carp mobile sensing framework in flutter. Framework, June 2019. Retrieved June 15, 2019 from `https://pub.dev/packages/carp_mobile_sensing`.

[14] Jakob E. Bardram and Mads Frost. The Personal Health Technology Design Space. *IEEE Pervasive Computing*, 15(2):70–78, 2016.

[15] Jakob E. Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. Designing mobile health technology for bipolar disorder: A field trial of the monarca system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2627–2636, New York, NY, USA, 2013. ACM.

[16] Jakob E. Bardram, Mads Frost, Nanna Tuxen, Maria Faurholt-Jepsen, and Lars V. Kessing. Designing context-aware cognitive behavioral therapy for unipolar and bipolar disorders. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, UbiComp '16, pages 1162–1170, New York, NY, USA, 2016. ACM.

[17] Jakob E Bardram and Mads M Frost. Double-loop health technology: enabling socio-technical design of personal health technology in clinical practice. In *Designing Healthcare That Works*, pages 167–186. Elsevier, 2018.

[18] Aaron T Beck. *Cognitive therapy of depression*. Guilford press, 1979.

[19] Dennis Becker, Vincent Bremer, Burkhardt Funk, Joost Asselbergs, Heleen Riper, and Jeroen Ruwaard. How to Predict Mood? Delving into Features of Smartphone-Based Data. In *Twenty-second Americas Conference on Information Systems*, pages 1–10, 2016.

[20] Till Beiwinkel, Sally Kindermann, Andreas Maier, Christopher Kerl, Jörn Moock, Guido Barbian, and Wulf Rössler. Using Smartphones to Monitor Bipolar Disorder Symptoms: A Pilot Study. *JMIR Mental Health*, 3(1):e2, 2016.

[21] Dror Ben-Zeev, Emily A Scherer, Rui Wang, Haiyi Xie, and Andrew T Campbell. Next-Generation Psychiatric Assessment: Using Smartphone Sensors to Monitor Behavior and Mental Health HHS Public Access. *Psychiatr Rehabil J*, 38(3):218–226, 2015.

[22] James Bennett-Levy, Paul Farrand, Helen Christensen, and Kathy Griffiths. *Oxford guide to low intensity CBT interventions*. Oxford University Press, 2010.

[23] Ethan M Berke, Tanzeem Choudhury, Shahid Ali, and Mashfiqui Rabbi. Objective measurement of sociability and activity: Mobile sensing in the community. *Annals of Family Medicine*, 9(4):344–350, 2011.

[24] Jan O. Berle, Erik R. Hauge, Ketil J. Oedegaard, Fred Holsten, and Ole B. Fasmer. Actigraphic registration of motor activity reveals a more structured behavioural pattern in schizophrenia than in major depression. *BMC Research Notes*, 3, 2010.

[25] Natalie Berry, Sandra Bucci, and Fiona Lobban. Use of the Internet and Mobile Phones for Self-Management of Severe Mental Health Problems: Qualitative Study of Staff Views. *JMIR mental health*, 4(4):e52, 2017.

[26] Dinesh Bhugra, Allan Tasman, Soumitra Pathare, Stefan Priebe, Shubulade Smith, John Torous, Melissa R Arbuckle, Alex Langford, Renato D Alarcón, Helen Fung Kum Chiu, Michael B First, Jerald Kay, Charlene Sunkel, Anita Thapar, Pichet Udomratn, Florence K Baingana, Dévora Kestel, Roger Man Kin Ng, Anita Patel, Livia De Picker, Kwame Julius McKenzie, Driss Moussaoui, Matt Muijen, Peter Bartlett, Sophie Davison, Tim Exworthy, Nasser Loza, Diana Rose, Julio Torales, Mark Brown, Helen Christensen, Joseph Firth, Matcheri Keshavan, Ang Li, Jukka-Pekka Onnela, Til Wykes, Hussien Elkholy, Gurvinder Kalra, Kate F Lovett, Michael J Travis, and Antonio Ventriglio. The WPA-Lancet Psychiatry Commission on the Future of Psychiatry. *The Lancet Psychiatry*, 4(10):775–818, 2017.

[27] Howard G Birnbaum, Ronald C Kessler, David Kelley, Rym Ben-Hamadi, Vijay N Joish, and Paul E Greenberg. Employer burden of mild, moderate, and severe major depressive disorder: mental health services utilization and costs, and work performance. *Depression and anxiety*, 27(1):78–89, 2010.

[28] R Clifford Blair and Walt Karniski. An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, 30(5):518–524, 1993.

[29] Ann Blandford, Jo Gibbs, Nikki Newhouse, Olga Perski, Aneesha Singh, and Elizabeth Murray. Seven lessons for interdisciplinary research on interactive digital health interventions. *Digital health*, 4:2055207618770325, 2018.

[30] Thomas Boillat, Homero Rivas, and Katarzyna Wac. "healthcare on a wrist": Increasing compliance through checklists on wearables in obesity (self-) management programs. In *Digital health*, pages 65–81. Springer, 2018.

[31] Catherine Briand, Juliette Sablier, Julie Anne Therrien, Karine Charbonneau, Jean François Pelletier, and Rhoda Weiss-Lambrou. Use of a mobile device in mental health rehabilitation: A clinical and comprehensive analysis of 11 cases. *Neuropsychological Rehabilitation*, 28(5):832–863, 2018.

[32] Marion Buchenau and Jane Fulton Suri. Experience prototyping. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 424–433. ACM, 2000.

[33] Nigel E Bush, Gary Ouellette, and Julie Kinn. Utility of the t2 mood tracker mobile application among army warrior transition unit service members. *Military medicine*, 179(12):1453–1457, 2014.

[34] Luca Canzian and Mirco Musolesi. Trajectories of Depression : Unobtrusive Monitoring of Depressive States by means of Smartphone Mobility Traces Analysis. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.*, pages 1293–1304, 2015.

[35] Denise M Casey. The historical development of distance education through technology. *TechTrends*, 52(2):45–51, 2008.

[36] Beenish M. Chaudry, Kay H. Connelly, Katie A. Siek, and Janet L. Welch. Mobile interface design for low-literacy populations. In *Proceedings of the 2Nd ACM SIGHIT International Health Informatics Symposium*, IHI '12, pages 91–100, New York, NY, USA, 2012. ACM.

[37] Helen Chen. Region south denmark clinics widely implement monsenso mhealth solution. Homepage, September 2017. Retrieved July 02, 2019 from `https://www.monsenso.com/region-south-denmark-clinics-implement-monsenso-mhealth-solution/`.

[38] Yong Min Cho, Hee Jin Lim, Hoon Jang, Kyunghee Kim, Jae Wook Choi, Chol Shin, Seung Ku Lee, Jong Hwa Kwon, and Nam Kim. A cross-sectional study of the association between mobile phone use and symptoms of ill health. *Environ Health Toxicol*, pages 1–7, 2016.

[39] Eun Kyoung Choe, Bongshin Lee, Matthew Kay, Wanda Pratt, and Julie A. Kientz. Sleeptight: Low-burden, self-monitoring technology for capturing and reflecting on sleep behaviors. In *Proceedings of the 2015 ACM International*

Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, pages 121–132, New York, NY, USA, 2015. ACM.

[40] Philip I Chow, Karl Fua, Yu Huang, Wesley Bonelli, Haoyi Xiong, Laura E Barnes, and Bethany A Teachman. Using Mobile Sensing to Test Clinical Models of Depression, Social Anxiety, State Affect, and Social Isolation Among College Students. *Journal of Medical Internet Research*, 19(3):e62, 2017.

[41] Victor P. Cornet and Richard J. Holden. Systematic review of smartphone-based passive sensing for health and wellbeing. *Journal of Biomedical Informatics*, 77(October 2017):120–132, 2018.

[42] William Coryell, William Scheftner, Martin Keller, Jean Endicott, Jack Maser, Gerald L Klerman, et al. The enduring psychosocial consequences of mania and depression. *American Journal of Psychiatry*, 150:720–720, 1993.

[43] Nick Craddock and Liz Forty. Genetics of affective (mood) disorders. *European Journal of Human Genetics*, 14(6):660, 2006.

[44] Pim Cuijpers, Annemieke van Straten, and Lisanne Warmerdam. Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review*, 27(3):318–326, 2007.

[45] Jennifer Dahne, Anahi Collado, C. W. Lejuez, Cristina M. Risco, Vanessa A. Diaz, Lisa Coles, Jacob Kustanowitz, Michael J. Zvolensky, and Matthew J. Carpenter. Pilot randomized controlled trial of a Spanish-language Behavioral Activation mobile app (¡Aptívate!) for the treatment of depressive symptoms among united states Latinx adults with limited English proficiency. *Journal of Affective Disorders*, 250(February):210–217, 2019.

[46] Mark Deady, David Johnston, David Milne, Nick Glozier, Dorian Peters, Rafael Calvo, and Samuel Harvey. Preliminary effectiveness of a smartphone app to reduce depressive symptoms in the workplace: Feasibility and acceptability study. *Journal of Medical Internet Research*, 20(12), 2018.

[47] Orianna Demasi, Adrian Aguilera, and Benjamin Recht. Detecting change in depressive symptoms from daily wellbeing questions, personality, and activity. In *2016 IEEE Wireless Health (WH)*, pages 1–8. IEEE, 2016.

[48] Anind K. Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin-Hyuk Hong, and Julian Ramos. Getting closer: An empirical investigation of the proximity of user to their smart phones. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, UbiComp '11, pages 163–172, New York, NY, USA, 2011. ACM.

[49] Robert F Dickerson, Eugenia I Gorlin, and John A Stankovic. Empath: a continuous remote emotional health monitoring system for depressive illness.

Proceedings of the 2nd Conference on Wireless Health - WH '11, page Art. 5, 2011.

[50] Sona Dimidjian, Steven D Hollon, Keith S Dobson, Karen B Schmaling, Robert J Kohlenberg, Michael E Addis, Robert Gallop, Joseph B McGlinchey, David K Markley, Jackie K Gollan, David C Atkins, David L Dunner, and Neil S Jacobson. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of consulting and clinical psychology*, 74(4):658–70, 2006.

[51] Keith S Dobson, Steven D Hollon, Sona Dimidjian, Karen B Schmaling, Robert J Kohlenberg, Robert J Gallop, Shireen L Rizvi, Jackie K Gollan, David L Dunner, and Neil S Jacobson. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *Journal of consulting and clinical psychology*, 76(3):468, 2008.

[52] Ezgi Dogan, Christian Sander, Xenija Wagner, Ulrich Hegerl, and Elisabeth Kohls. Smartphone-Based Monitoring of Objective and Subjective Data in Affective Disorders: Where Are We and Where Are We Going? Systematic Review. *Journal of Medical Internet Research*, 19(7):e262, 2017.

[53] Gavin Doherty, David Coyle, and Mark Matthews. Design and evaluation guidelines for mental health technologies. *Interacting with Computers*, 22(4):243–252, 2010.

[54] Kevin Doherty and Gavin Doherty. Engagement in hci: Conception, theory and measurement. *ACM Comput. Surv.*, 51(5):99:1–99:39, November 2018.

[55] Kevin Doherty, José Marcano-Belisario, Martin Cohn, Nikolaos Mastellos, Cecily Morrison, Josip Car, and Gavin Doherty. Engagement with mental health screening on mobile devices: Results from an antenatal feasibility study. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 186:1–186:15, New York, NY, USA, 2019. ACM.

[56] Afsaneh Doryab, Jun Ki Min, Jason Wiese, John Zimmerman, and Jason I Hong. Detection of behavior change in people with depression. In *AAAI Workshops Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 12–16, 2014.

[57] Meghan K. Edwards and Paul D. Loprinzi. Effects of a Sedentary Behavior???Inducing Randomized Controlled Intervention on Depression and Mood Profile in Active Young Adults. *Mayo Clinic Proceedings*, 91(8):984–998, 2016.

[58] David Ekers, Lisa Webster, Annemieke Van Straten, Pim Cuijpers, David Richards, and Simon Gilbody. Behavioural activation for depression; An update of meta-analysis of effectiveness and sub group analysis. *PLoS ONE*, 9(6), 2014.

[59] Michael D Ekstrand, John T Riedl, Joseph A Konstan, et al. Collaborative filtering recommender systems. *Foundations and Trends® in Human–Computer Interaction*, 4(2):81–173, 2011.

[60] Farhan, Yue, Morillo, Ware, Lu, Bi, Kamath, Russell, Bamis, and Wang. Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data. In *2016 IEEE Wireless Health (WH)*, pages 1–8. IEEE, oct 2016.

[61] M Faurholt-Jepsen, J Busk, M Frost, M Vinberg, E M Christensen, O Winther, J E Bardram, and L V Kessing. Voice analysis as an objective state marker in bipolar disorder. *Translational Psychiatry*, 6(7):e856, 2016.

[62] Maria Faurholt-jepsen, Søren Brage, Maj Vinberg, Ellen Margrethe, Ulla Knorr, Hans Mørch, and Lars Vedel. Differences in psychomotor activity in patients suffering from unipolar and bipolar affective disorder in the remitted or mild / moderate depressive state. *Journal of Affective Disorders*, 141(2-3):457–463, 2012.

[63] Maria Faurholt-Jepsen, Mads Frost, Maj Vinberg, Ellen Margrethe Christensen, Jakob E. Bardram, and Lars Vedel Kessing. Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Research*, 217(1-2):124–127, 2014.

[64] Maria Faurholt-Jepsen, Maj Vinberg, Mads Frost, Ellen Margrethe Christensen, Jakob E. Bardram, and Lars Vedel Kessing. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. *Bipolar Disorders*, 17(7):715–728, 2015.

[65] Maria Faurholt-Jepsen, Maj Vinberg, Mads Frost, Sune Debel, Ellen Margrethe Christensen, Jakob E. Bardram, and Lars Vedel Kessing. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. *International Journal of Methods in Psychiatric Research*, 25(4):309–323, 2016.

[66] James M Ferguson. Ssri antidepressant medications: adverse effects and tolerability. *Primary care companion to the Journal of clinical psychiatry*, 3(1):22, 2001.

[67] Ben Fletcher and Karen J Pine. *Flex: Do something different.* Univ of Hertfordshire Press, 2012.

[68] Forbes. Ten most innovative companies in europe 2015. `https://www.forbes.com/pictures/55cbb457e4b0ffa7afe3d731/4-coloplast`, August 2015.

[69] Matthew Fuller-Tyszkiewicz, Ben Richardson, Britt Klein, Helen Skouteris, Helen Christensen, David Austin, David Castle, Cathrine Mihalopoulos, Renee O'Donnell, Lilani Arulkadacham, Adrian Shatte, and Anna Ware. A mobile

app-based intervention for depression: End-user and expert usability testing study. *Journal of Medical Internet Research*, 20(8):1–12, 2018.

[70] Shiv Gautam, Akhilesh Jain, Manaswi Gautam, Vihang N Vahia, and Sandeep Grover. Clinical practice guidelines for the management of depression. *Indian journal of psychiatry*, 59(Suppl 1):S34, 2017.

[71] Linda K George, Dan G Blazer, Dana C Hughes, and Nancy Fowler. Social support and the outcome of major depression. *The British Journal of Psychiatry*, 154(4):478–485, 1989.

[72] Anda Gershon, Nilam Ram, Sheri L. Johnson, Allison G. Harvey, and Jamie M. Zeitzer. Daily actigraphy profiles distinguish depressive and interepisode states in bipolar disorder. *Clinical Psychological Science*, 4(4):641–650, 2016.

[73] Tasha Glenn and Scott Monteith. New Measures of Mental State and Behavior Based on Data Collected From Sensors, Smartphones, and the Internet. *Current Psychiatry Reports*, 16(12), 2014.

[74] Kimberly Goldapple, Zindel Segal, Carol Garson, Mark Lau, Peter Bieling, Sidney Kennedy, and Helen Mayberg. Modulation of cortical-limbic pathways in major depression: treatment-specific effects of cognitive behavior therapy. *Archives of general psychiatry*, 61(1):34–41, 2004.

[75] Robert Gonzalez, Carol A. Tamminga, Mauriclo Tohen, and Trisha Suppes. The relationship between affective state and the rhythmicity of activity in bipolar disorder. *Journal of Clinical Psychiatry*, 75(4):1–14, 2014.

[76] David Gotz and David Borland. Data-Driven Healthcare: Challenges and Opportunities for Interactive Visualization. *IEEE Computer Graphics and Applications*, 36(3):90–96, 2016.

[77] Franz Gravenhorst, Amir Muaremi, Jakob Bardram, Agnes Grünerbl, Oscar Mayora, Gabriel Wurzer, Mads Frost, Venet Osmani, Bert Arnrich, Paul Lukowicz, and Others. Mobile phones as medical devices in mental disorder treatment: an overview. *Personal and Ubiquitous Computing*, 19(2):335–353, 2015.

[78] Gerry Greenstone. The history of bloodletting. *BC Medical Journal*, 52(1):12–14, 2010.

[79] Rebecca Grist, Joanna Porter, and Paul Stallard. Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. *Journal of medical Internet research*, 19(5):e176, 2017.

[80] Agnes Grünerbl, Amir Muaremi, Venet Osmani, Gernot Bahle, Stefan Öhler, Gerard Tröster, Oscar Mayora, Christian Haring, and Paul Lukowicz. Smart-Phone Based Recognition of States and State Changes in Bipolar Disorder Patients. *IEEE Journal of Biomedical and Health Informatics*, 19(1):140–148, 2015.

[81] A Guidi, N Vanello, G Bertschy, C Gentili, L Landini, and E P Scilingo. Biomedical Signal Processing and Control Automatic analysis of speech F0 contour for the characterization of mood changes in bipolar patients. *Biomedical Signal Processing and Control*, 17:29–37, 2015.

[82] E. E. Haroz, M. Ritchey, J. K. Bass, B. A. Kohrt, J. Augustinavicius, L. Michalopoulos, M. D. Burkey, and P. Bolton. How is depression experienced around the world? A systematic review of qualitative literature. *Social Science and Medicine*, 183:151–162, 2017.

[83] Ainslie Hatch, Julia E. Hoffman, Ruth Ross, and John P. Docherty. Expert consensus survey on digital health tools for patients with serious mental illness: Optimizing for user characteristics and user support. *Journal of Medical Internet Research*, 20(6), 2018.

[84] Erik R. Hauge, Jan Øystein Berle, Ketil J. Oedegaard, Fred Holsten, and Ole Bernt Fasmer. Nonlinear analysis of motor activity shows differences between schizophrenia and depression: A study using fourier analysis and sample entropy. *PLoS ONE*, 6(1):1–10, 2011.

[85] Philip Henson, Gary David, Karen Albright, and John Torous. Deriving a practical framework for the evaluation of health apps. *The Lancet Digital Health*, 1(2):e52–e54, 2019.

[86] Stefan G Hofmann, Anu Asnaani, Imke JJ Vonk, Alice T Sawyer, and Angela Fang. The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive therapy and research*, 36(5):427–440, 2012.

[87] Victoria Hollis, Artie Konrad, Aaron Springer, Matthew Antoun, Christopher Antoun, Rob Martin, and Steve Whittaker. What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being. *Human-Computer Interaction*, 32(5-6):208–267, 2017.

[88] Steven D Hollon, Robert J DeRubeis, Richard C Shelton, Jay D Amsterdam, Ronald M Salomon, John P O'Reardon, Margaret L Lovett, Paula R Young, Kirsten L Haman, Brent B Freeman, and Robert Gallop. Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. *Archives of general psychiatry*, 62(4):417–22, 2005.

[89] Kasper Hornbaek. Some Whys and Hows of Experiments in Human–Computer Interaction. *Foundations and Trends® in Human–Computer Interaction*, 5(4):299–373, 2011.

[90] Stephanie Houde and Charles Hill. What do prototypes prototype? In *Handbook of human-computer interaction*, pages 367–381. Elsevier, 1997.

[91] Anna Huguet, Sanjay Rao, Patrick J. McGrath, Lori Wozney, Mike Wheaton, Jill Conrod, and Sharlene Rozario. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS ONE*, 11(5):1–19, 2016.

[92] NS Jacobson, KS Dobson, PA Truax, ME Addis, K Koerner, JK Gollan, E Gortner, and SE Prince. A component analysis of cognitive-behavioral treatment for depression. *Journal of consulting and clinical psychology*, 64(2):295–304, 1996.

[93] A John Rush, Madhukar H Trivedi, Stephen R Wisniewski, Andrew A Nierenberg, Jonathan W Stewart, Diane Warden, Mba George Niederehe, Michael E Thase, Philip W Lavori, Barry D Lebowitz, Patrick J McGrath, Jerrold F Rosenbaum, Harold A Sackeim, David J Kupfer, James Luther, and Ma Maurizio Fava. STAR-D (2006; AjPsych) Tiered approach for depression. *Am J Psychiatry*, 16311(November):1905–1917, 2006.

[94] Michael I Jordan and Tom M Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.

[95] Laura Aller Jónasdóttir and Sara Pagh Kofoed. Patienter føler sig i ingenmandsland. Homepage, August 2019. Retrieved August 28, 2019 from `https://www.dr.dk/nyheder/indland/svaere-overgange-i-psykiatrien-patienter-foeler-sig-i-ingenmandsland`.

[96] Andrew B Kahng. Scaling: More than moore's law. *IEEE Design & Test of Computers*, 27(3):86–87, 2010.

[97] Sylvia Deidre Kauer, Sophie Caroline Reid, Alexander Hew Dale Crooke, Angela Khor, Stephen John Charles Hearps, Anthony Francis Jorm, Lena Sanci, and George Patton. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. *Journal of medical Internet research*, 14(3):e67, 2012.

[98] Christina Kelley, Bongshin Lee, and Lauren Wilcox. Self-tracking for Mental Wellness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pages 629–641, 2017.

[99] Ronald C Kessler. The costs of depression. *Psychiatric Clinics*, 35(1):1–14, 2012.

[100] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. How to evaluate technologies for health behavior change in hci research. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 3063–3072. ACM, 2011.

[101] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. How to evaluate technologies for health behavior change in hci research. In *Proceedings of the SIGCHI*

Conference on Human Factors in Computing Systems, CHI '11, pages 3063–3072, New York, NY, USA, 2011. ACM.

[102] Robert Kohn, Shekhar Saxena, Itzhak Levav, and Benedetto Saraceno. The treatment gap in mental health care. *Bulletin of the World Health Organization*, 82(11):858–866, 2004.

[103] Karoline Krane-Gartiser, Tone Elise Gjotterud Henriksen, Gunnar Morken, Arne Vaaler, and Ole Bernt Fasmer. Actigraphic assessment of motor activity in acutely admitted inpatients with bipolar disorder. *PLoS ONE*, 9(2), 2014.

[104] C. W. Lejuez, Derek R. Hopko, Ron Acierno, Stacey B. Daughters, and Sherry L. Pagoto. Ten year revision of the brief behavioral activation treatment for depression: Revised treatment manual. *Behavior Modification*, 35(2):111–161, 2011.

[105] C W. Lejuez, Derek R.; Hopko, and Sandra D Hopko. A brief behavioral activation treatment for depression: Treatment manual. *Behavior Modification*, 25(2):255–286, 2001.

[106] Andrew C Leon, Lori L Davis, and Helena C Kraemer. The role and interpretation of pilot studies in clinical research. *Journal of psychiatric research*, 45(5):626–629, 2011.

[107] Peter M Lewinsohn. A behavioral approach to depression. *Essential papers on depression*, pages 150–72, 1974.

[108] James R Lewis. Psychometric evaluation of the post-study system usability questionnaire: The pssuq. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 36, pages 1259–1260. SAGE Publications Sage CA: Los Angeles, CA, 1992.

[109] Mu Lin, Nicholas D. Lane, Mashfiqui Mohammod, Xiaochao Yang, Hong Lu, Giuseppe Cardone, Shahid Ali, Afsaneh Doryab, Ethan Berke, Andrew T. Campbell, and Tanzeem Choudhury. Bewell+: Multi-dimensional wellbeing monitoring with community-guided user feedback and energy optimization. In *Proceedings of the Conference on Wireless Health*, WH '12, pages 10:1–10:8, New York, NY, USA, 2012. ACM.

[110] Klaus Linde, Gerta Rücker, Kirsten Sigterman, Susanne Jamil, Karin Meissner, Antonius Schneider, and Levente Kriston. Comparative effectiveness of psychological treatments for depressive disorders in primary care: network meta-analysis. *BMC family practice*, 16(1):103, 2015.

[111] Jessica Lipschitz, Christopher J Miller, Timothy P Hogan, Katherine E Burdick, Rachel Lippin-Foster, Steven R Simon, and James Burgess. Adoption of Mobile Apps for Depression and Anxiety: Cross-Sectional Survey Study on Patient Interest and Barriers to Engagement. *JMIR Mental Health*, 6(1):e11334, 2019.

[112] Paul D Loprinzi and Sara Mahoney. Concurrent occurrence of multiple positive lifestyle behaviors and depression among adults in the United States. *Journal of Affective Disorders*, 165:126–130, 2014.

[113] Kien Hoa Ly, Elsa Janni, Richard Wrede, Mina Sedem, Tara Donker, Per Carlbring, and Gerhard Andersson. Experiences of a guided smartphone-based behavioral activation therapy for depression: A qualitative study. *Internet Interventions*, 2(1):60–68, 2015.

[114] Kien Hoa Ly, Naira Topooco, Hanna Cederlund, Anna Wallin, Jan Bergstrom, Olof Molander, Per Carlbring, and Gerhard Andersson. Smartphone-supported versus full behavioural activation for depression: A randomised controlled trial. *PLoS ONE*, 10(5):1–16, 2015.

[115] Kien Hoa Ly, Anna Trüschel, Linnea Jarl, Susanna Magnusson, Tove Windahl, Robert Johansson, Per Carlbring, and Gerhard Andersson. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. *BMJ open*, 4(1):e003440, 2014.

[116] Wendy E Mackay and Anne-Laure Fayard. Hci, natural science and design: a framework for triangulation across disciplines. In *Symposium on Designing Interactive Systems: Proceedings of the 2 nd conference on Designing interactive systems: processes, practices, methods, and techniques*, volume 18, pages 223–234, 1997.

[117] I Scott MacKenzie. *Human-computer interaction: An empirical research perspective*. Newnes, 2012.

[118] DJ MacPhillamy and PM Lewinsohn. The measurement of reinforcing events. In *Proceedings of the 80th Annual Convention of the American Psychological Association*, volume 7, pages 399–400, 1972.

[119] Douglas J. MacPhillamy and Peter M. Lewinsohn. The pleasant events schedule: Studies on reliability, validity, and scale intercorrelation. *Journal of Consulting and Clinical Psychology*, 50(3):363–380, 1982.

[120] Gabriela Marcu, Jakob E Bardram, and Silvia Gabrielli. A framework for overcoming challenges in designing persuasive monitoring and feedback systems for mental illness. *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, pages 1–8, 2011.

[121] Gloria Mark, Mary Czerwinski, Shamsi Iqbal, and Paul Johns. Workplace Indicators of Mood: Behavioral and Cognitive Correlates of Mood Among Information Workers. *Proceedings of the 6th International Conference on Digital Health Conference*, pages 29–36, 2016.

[122] Ruth Martell, Christopher R and Dimidjian, Sona and Herman-Dunn. *Behavioral Activation for Depression: A Clinician's Guide*, volume 40. Guilford Press, mar 2013.

[123] Colin D Mathers and Dejan Loncar. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine*, 3(11):e442, 2006.

[124] Aleksandar Matic, Venet Osmani, Andrei Popleteev, and Oscar Mayora-Ibarra. Smart phone sensing to examine effects of social interactions and non-sedentary work time on mood changes. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6967 LNAI:200–213, 2011.

[125] Kibby McMahon, Nathaniel R Herr, Noga Zerubavel, Nicolas Hoertel, and Andrada D Neacsiu. Psychotherapeutic treatment of bipolar depression. *Psychiatric Clinics*, 39(1):35–56, 2016.

[126] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. Towards multi-modal anticipatory monitoring of depressive states through the analysis of human-smartphone interaction. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16*, pages 1132–1138, 2016.

[127] Helena M. Mentis, Anita Komlodi, Katrina Schrader, Michael Phipps, Ann Gruber-Baldini, Karen Yarbrough, and Lisa Shulman. Crafting a View of Self-Tracking Data in the Clinical Visit. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pages 5800–5812, 2017.

[128] Madhura Mestry, Jigesh Mehta, Ajit Mishra, and Kiran Gawande. Identifying associations between smartphone usage and mental health during depression, anxiety and stress. *Proceedings - 2015 International Conference on Communication, Information and Computing Technology, ICCICT 2015*, 2015.

[129] Hiroyasu Miwa, Shin-ichiro Sasahara, and Toshihiro Matsui. Roll-over detection and sleep quality measurement using a wearable sensor. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2007:1507–1510, aug 2007.

[130] David Moher, Alessandro Liberati, Jennifer Tetzlaff, D G Altman, and Prisma Grp. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Physical Therapy*, 89(9):873–880, jul 2009.

[131] David C Mohr, Joyce Ho, Jenna Duffecy, Kelly G Baron, Kenneth A Lehman, Ling Jin, and Douglas Reifler. Perceived barriers to psychological treatments and their relationship to depression. *Journal of clinical psychology*, 66(4):394–409, 2010.

[132] David C. Mohr, Ken R. Weingardt, Madhu Reddy, and Stephen M. Schueller. Three Problems With Current Digital Mental Health Research . . . and Three Things We Can Do About Them. *Psychiatric Services*, 68(5):427–429, 2017.

[133] David C. Mohr, Mi Zhang, and Stephen M. Schueller. Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning. *Annual Review of Clinical Psychology*, 13(1):23–47, may 2017.

[134] Paul Monroe. *A Text-book in the History of Education*. New York: Macmillan, 1905.

[135] Merete M Mørch and Nicole K Rosenberg. *Kognitiv terapi: modeller og metoder*. Gyldendal A/S, 2005.

[136] Amir Muaremi, Franz Gravenhorst, Agnes Grünerbl, Bert Arnrich, and Gerhard Tröster. Assessing bipolar episodes using speech cues derived from phone calls. In *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, volume 100, pages 103–114, 2014.

[137] Alan F. Newell and Peter Gregor. “user sensitive inclusive design”— in search of a new paradigm. In *Proceedings on the 2000 Conference on Universal Usability*, CUU '00, pages 39–44, New York, NY, USA, 2000. ACM.

[138] Jennifer Nicholas, Mark Erik Larsen, Judith Proudfoot, and Helen Christensen. Mobile apps for bipolar disorder: a systematic review of features and content quality. *Journal of medical Internet research*, 17(8):e198, 2015.

[139] Jakob Nielsen. *Usability engineering*. Elsevier, 1994.

[140] Jakob Nielsen. 10 usability heuristics for user interface design. *Nielsen Norman Group*, 1(1), 1995.

[141] Susan Nolen-Hoeksema. Responses to depression and their effects on the duration of depressive episodes. *Journal of abnormal psychology*, 100(4):569, 1991.

[142] Don Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[143] A Novotney. Should you use an app to help that client. *Monitor Psychol*, 47:64–66, 2016.

[144] J. T. O'Brien, P. Gallagher, D. Stow, N. Hammerla, T. Ploetz, M. Firbank, C. Ladha, K. Ladha, D. Jackson, R. McNaney, I. N. Ferrier, and P. Olivier. A study of wrist-worn activity measurement as a potential real-world biomarker for late-life depression. *Psychological Medicine*, 47(1):93–102, 2017.

[145] US Bureau of Labor Statistics. American time use survey - 2014 results. pdf, June 2014. Retrieved July 23, 2019 from `https://www.bls.gov/news.release/archives/atus_06242015.pdf`.

[146] Miranda Olff. Mobile mental health: a challenging research agenda. *European journal of psychotraumatology*, 6(1):27882, 2015.

[147] World Health Organization. Suicide. Homepage, August 2018. Retrieved June 21, 2019 from `https://www.who.int/mental_health/suicide-prevention`.

[148] Venet Osmani, Alban Maxhuni, Agnes Grünerbl, Paul Lukowicz, Christian Haring, and Oscar Mayora. Monitoring activity of patients with bipolar disorder using smart phones. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia - MoMM '13*, pages 85–92, 2013.

[149] Christian Otte, Stefan M Gold, Brenda W Penninx, Carmine M Pariante, Amit Etkin, Maurizio Fava, David C Mohr, and Alan F Schatzberg. Major depressive disorder. *Nature reviews Disease primers*, 2:16065, 2016.

[150] N. Palmius, A. Tsanas, K. E. A. Saunders, A. C. Bilderbeck, J. R. Geddes, G. M. Goodwin, and M. De Vos. Detecting Bipolar Depression From Geographic Location Data. *IEEE Transactions on Biomedical Engineering*, 64(8):1761–1771, 2017.

[151] Adam Pedley. How flutter works. Homepage, April 2018. Retrieved July 21, 2019 from `https://buildflutter.com/how-flutter-works/`.

[152] David Peiris, J. Jaime Miranda, and David C. Mohr. Going beyond killer apps: Building a better mHealth evidence base. *BMJ Global Health*, 3(1):2017–2019, 2018.

[153] Brenda WJH Penninx, Yuri Milaneschi, Femke Lamers, and Nicole Vogelzangs. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC medicine*, 11(1):129, 2013.

[154] Charles Petzold. *Creating Mobile Apps with Xamarin. Forms Preview Edition 2*. Microsoft Press, 2015.

[155] Vivek Pillai, Lindsey A. Steenburg, Jeffrey A. Ciesla, Thomas Roth, and Christopher L. Drake. A seven day actigraphy-based study of rumination and sleep disturbance among young adults with depressive symptoms. *Journal of Psychosomatic Research*, 77(1):70–75, 2014.

[156] Frederic M Quitkin, Patrick J McGrath, et al. Different types of placebo response in patients receiving antidepressants. *The American journal of psychiatry*, 148(2):197, 1991.

[157] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang, and Tanzeem Choudhury. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. *Ubicomp '15, September 7–11, 2015, Osaka, Japan*, pages 707–718, 2015.

[158] R. Rana, M. Hume, J. Reilly, R. Jurdak, and J. Soar. Opportunistic and Context-aware Affect Sensing on Smartphones: The Concept, Challenges and Opportunities. *arXiv e-prints*, February 2015.

[159] Stefan Rennick-Egglestone, Sarah Knowles, Gill Toms, Penny Bee, Karina Lovell, and Peter Bower. Health Technologies 'In the Wild': Experiences of Engagement with Computerised CBT. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2124–2135, 2016.

[160] David A Richards, David Ekers, Dean McMillan, Rod S Taylor, Sarah Byford, Fiona C Warren, Barbara Barrett, Paul A Farrand, Simon Gilbody, Willem Kuyken, et al. Cost and outcome of behavioural activation versus cognitive behavioural therapy for depression (cobra): a randomised, controlled, non-inferiority trial. *The Lancet*, 388(10047):871–880, 2016.

[161] W David Robinson, Jenenne A Geske, Layne A Prest, and Rachel Barnacle. Depression treatment in primary care. *J Am Board Fam Pract*, 18(2):79–86, mar 2005.

[162] John Rooksby, Alistair Morrison, and Dave Murray-Rust. Student perspectives on digital phenotyping: The acceptability of using smartphone data to assess mental health. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 425. ACM, 2019.

[163] Sohrab Saeb, Emily G. Lattie, Stephen M. Schueller, Konrad P. Kording, and David C. Mohr. The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ*, 4:e2537, 2016.

[164] Sohrab Saeb, Mi Zhang, Christopher J. Karr, Stephen M. Schueller, Marya E. Corden, Konrad P. Kording, and David C. Mohr. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *Journal of Medical Internet Research*, 17(7):1–11, 2015.

[165] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, and Corina Sas. Hci and affective health: Taking stock of a decade of studies and charting future research directions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 245:1–245:17, New York, NY, USA, 2019. ACM.

[166] K.E.A. Saunders, A.C. Bilderbeck, P. Panchal, L.Z. Atkinson, J.R. Geddes, and G.M. Goodwin. Experiences of remote mood and activity monitoring in bipolar disorder: A qualitative study. *European Psychiatry*, 41:115–121, 2017.

[167] Connie Schardt, Martha B Adams, Thomas Owens, Sheri Keitz, and Paul Fontelo. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC medical informatics and decision making*, 7:16, 2007.

[168] Danielle A Schlosser, Timothy R Campellone, Brandy Truong, Kevin Etter, Silvia Vergani, Kiya Komaiko, and Sophia Vinogradov. Efficacy of PRIME , a Mobile App Intervention Designed to Improve Motivation in Young People With Schizophrenia. *Schizophrenia bulletin*, 44(5):1010–1020, 2018.

[169] Jan Scott, Greg Murray, Chantal Henry, Gunnar Morken, Elizabeth Scott, Jules Angst, Kathleen R. Merikangas, and Ian B. Hickie. Activation in bipolar disorders a systematic review. *JAMA Psychiatry*, 74(2):189–196, 2017.

[170] Sandra Servia-Rodríguez, Kiran K. Rachuri, Cecilia Mascolo, Peter J. Rentfrow, Neal Lathia, and Gillian M. Sandstrom. Mobile sensing at the service of mental well-being: A large-scale longitudinal study. In *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, pages 103–112, Republic and Canton of Geneva, Switzerland, 2017. International World Wide Web Conferences Steering Committee.

[171] Adrian B.R. Shatte, Delyse M. Hutchinson, and Samantha J. Teague. Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, 2019.

[172] Ryan J. Shaw, Dori M. Steinberg, Jonathan Bonnet, Farhad Modarai, Aaron George, Traven Cunningham, Markedia Mason, Mohammad Shahsahebi, Steven C. Grambow, Gary G. Bennett, and Hayden B. Bosworth. Mobile health devices: Will patients actually use them? *Journal of the American Medical Informatics Association*, 23(3):462–466, 2016.

[173] Mike Slade. Mental illness and well-being: the central importance of positive psychology and recovery approaches. *BMC health services research*, 10(1):26, 2010.

[174] John Snow. Cholera and the water supply in the south districts of london in 1854. *Journal of public health, and sanitary review*, 2(7):239, 1856.

[175] Daniel J Solove. A taxonomy of privacy. *U. Pa. L. Rev.*, 154:477, 2005.

[176] Isabelle Soucy Chartier and Martin D. Provencher. Behavioural activation for depression: Efficacy, effectiveness and dissemination. *Journal of Affective Disorders*, 145(3):292–299, 2013.

[177] Aaron Springer, Victoria Hollis, and Steve Whittaker. Mood modeling: accuracy depends on active logging and reflection. *Personal and Ubiquitous Computing*, 22(4):723–737, 2018.

[178] Julie St-Amand, Martin D. Provencher, Lynda Bélanger, and Charles M. Morin. Sleep disturbances in bipolar disorder during remission. *Journal of Affective Disorders*, 146(1):112–119, 2013.

[179] Victor J Strecher, Gerard H Seijts, Gerjo J Kok, Gary P Latham, Russell Glasgow, Brenda DeVellis, Ree M Meertens, and David W Bulger. Goal setting as a strategy for health behavior change. *Health education quarterly*, 22(2):190–200, 1995.

[180] Megan Teychenne, Kylie Ball, and Jo Salmon. Physical activity and likelihood of depression in adults: a review. *Preventive medicine*, 46(5):397–411, 2008.

[181] Nickolai Titov, Blake F. Dear, Lauren G. Staples, James Bennett-Levy, Britt Klein, Ronald M. Rapee, Clare Shann, David Richards, Gerhard Andersson, Lee Ritterband, Carol Purtell, Greg Bezuidenhout, Luke Johnston, and Olav B. Nielssen. MindSpot Clinic: An Accessible, Efficient, and Effective Online Treatment Service for Anxiety and Depression. *Psychiatric Services*, 66(10):1043–1050, 2015.

[182] Doron Todder, Serdal Caliskan, and Bernhard T. Baune. Longitudinal changes of day-time and night-time gross motor activity in clinical responders and non-responders of major depression. *World Journal of Biological Psychiatry*, 10(4):276–284, 2009.

[183] John Torous, Gerhard Andersson, Andrew Bertagnoli, Helen Christensen, Pim Cuijpers, Joseph Firth, Adam Haim, Honor Hsin, Chris Hollis, Shôn Lewis, David C. Mohr, Abhishek Pratap, Spencer Roux, Joel Sherrill, and Patricia A. Arean. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry*, 18(1):97–98, 2019.

[184] John Torous, Rohn Friedman, and Matcheri Keshvan. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *Journal of Medical Internet Research*, 16(1):1–8, 2014.

[185] John Torous, Jennifer Nicholas, Mark E Larsen, Joseph Firth, and Helen Christensen. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evidence-based mental health*, 21(3):116–119, 2018.

[186] John Torous and Laura Weiss Roberts. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. *JAMA psychiatry*, 74(5):437–438, 2017.

[187] John Torous, Patrick Staples, and Jukka Pekka Onnela. Realizing the Potential of Mobile Mental Health: New Methods for New Data in Psychiatry. *Current Psychiatry Reports*, 17(8), 2015.

[188] John Torous, Hannah Wisniewski, Gang Liu, and Matcheri Keshavan. Mental Health Mobile Phone App Usage, Concerns, and Benefits Among Psychiatric Outpatients: Comparative Survey Study. *JMIR Mental Health*, 5(4):e11715, 2018.

[189] John Torous, Jessica Woodyatt, Matcheri Keshavan, and Laura M. Tully. A new hope for early psychosis care: The evolving landscape of digital care tools. *British Journal of Psychiatry*, 214(5):269–272, 2019.

[190] Femina P Varghese and Gary R VandenBos. Clinical apps and public service settings: Innovations, diverse applications, and limits. *Psychological services*, 16(2):183, 2019.

[191] Giovanna N. Vilaza and Jakob E. Bardram. Sharing access to behavioural and personal health data: Designers' perspectives on opportunities and barriers. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth'19, pages 346–350, New York, NY, USA, 2019. ACM.

[192] Katarzyna Wac, Selim Ickin, Jin-Hyuk Hong, Lucjan Janowski, Markus Fiedler, and Anind K. Dey. Studying the experience of mobile applications used in different contexts of daily life. In *Proceedings of the First ACM SIGCOMM Workshop on Measurements Up the Stack*, W-MUST '11, pages 7–12, New York, NY, USA, 2011. ACM.

[193] Fabian Wahle, Tobias Kowatsch, Elgar Fleisch, Michael Rufer, and Steffi Weidt. Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild. *JMIR mHealth and uHealth*, 4(3):e111, 2016.

[194] Rui Wang, Andrew T. Campbell, and Xia Zhou. Using opportunistic face logging from smartphone to infer mental health: Challenges and future directions. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, UbiComp/ISWC'15 Adjunct, pages 683–692, New York, NY, USA, 2015. ACM.

[195] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 3–14, 2014.

[196] Peter West, Max Van Kleek, Richard Giordano, Mark J. Weal, and Nigel Shadbolt. Common Barriers to the Use of Patient-Generated Data Across Clinical Settings. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pages 1–13, 2018.

[197] Haoyi Xiong, Yu Huang, Laura E. Barnes, and Matthew S. Gerber. Sensus: A Cross-platform, General-purpose System for Mobile Crowdsensing in Human-subject Studies. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 415–426, 2016.

[198] Shiqi Yang, Ping Zhou, Kui Duan, M. Shamim Hossain, and Mohammed F. Alhamid. emHealth: Towards Emotion Health Through Depression Prediction and Intelligent Health Recommender System. *Mobile Networks and Applications*, 23(2):216–226, 2018.

[199] Jeffrey K Zeig and W Munion. *What is psychotherapy?: Contemporary perspectives.* Jossey-Bass, 1990.

[200] Xiaoyi Zhang, Laura R. Pina, and James Fogarty. Examining unlock journaling with diaries and reminders for in situ self-report in health and wellness. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5658–5664, New York, NY, USA, 2016. ACM.

[201] Haining Zhu, Yuhan Luo, and Eun Kyoung Choe. Making space for the quality care: Opportunities for technology in cognitive behavioral therapy for insomnia. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5773–5786, New York, NY, USA, 2017. ACM.