

# Benefits of Using Activity Recommender Technology for Self-management of Depressive Symptoms

DARIUS A. ROHANI, Department of Health Technology, Technical University of Denmark, Denmark  
MARIA FAURHOLT-JEPSEN and LARS V. KESSING, Psychiatric Centre Copenhagen, Denmark  
JAKOB E. BARDRAM, Department of Health Technology, Technical University of Denmark, Denmark

Behavioral Activation (BA) therapy has shown to be effective in treating depression. Recommending healthy activities is a core principle in BA, which is typically done by the therapist. However, most BA smartphone applications do not recommend specific activities. This article reports quantitative results from an 8-week feasibility study of a previously presented smartphone-based BA recommender system. The system supports the planning and enacting of pleasurable activities and promotes activation of diverse activity types. Enrollment included 43 clinically depressed patients who installed the system on their phone and initiated activity scheduling. Twenty-nine patients used the system daily for more than a week. These patients presented a significant reduction in depressive symptoms during the study period. They displayed a more personalized usage approach and created recurring health goals comprising of their own customized activities. Furthermore, they took inspiration within various types of activities, thereby displaying more activity diversity. This study suggests that enacting a diverse mixture of activities that promote good sleep, personal hygiene, exercise, social contact, and leisure time can be essential in managing depressive symptoms. A smartphone-based activity recommender system can help patients achieve this.

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

Additional Key Words and Phrases: Activity, depression, depressive symptoms, well-being, behavioral activation, smartphone, diversity, personalization, quantitative

## ACM Reference format:

Darius A. Rohani, Maria Faurholt-Jepsen, Lars V. Kessing, and Jakob E. Bardram. 2021. Benefits of Using Activity Recommender Technology for Self-management of Depressive Symptoms. *ACM Trans. Comput. Healthcare* 2, 4, Article 31 (August 2021), 21 pages.

<https://doi.org/10.1145/3462212>

This work is supported by the Copenhagen Center for Health Technology (CACHET) and the Innovation Fund Denmark under the project: "RADMIS: Reducing the rate and duration of readmission among patients with unipolar and bipolar disorder using smartphone-based monitoring and treatment" (<https://www.cachet.dk/research/finalized-research-projects/radmis>).

Authors' addresses: D. A. Rohani and J. E. Bardram, Department of Health Technology, Technical University of Denmark, Denmark; emails: [daroh@dtu.dk](mailto:daroh@dtu.dk), [jakba@dtu.dk](mailto:jakba@dtu.dk); M. Faurholt-Jepsen and L. V. Kessing, Psychiatric Centre Copenhagen, Denmark; emails: [maria.faurholtjepsen@regionh.dk](mailto:maria.faurholtjepsen@regionh.dk), [lars.vedel.kessing@regionh.dk](mailto:lars.vedel.kessing@regionh.dk).



This work is licensed under a Creative Commons Attribution-NonCommercial International 4.0 License.

© 2021 Association for Computing Machinery.

2637-8051/2021/08-ART31 \$15.00

<https://doi.org/10.1145/3462212>

## 1 INTRODUCTION

Inactiveness, lack of motivation, difficulties in decision-making, and diminished interest, pleasure, and energy to engage in activities are all symptoms observed in depressive disorders [2]. Depressive disorders are characterized by at least 2 weeks of symptoms in what is known as a depressive episode. It is estimated that one-fifth of the world population will experience at least one episode [16]. Situated in a depressive episode has severe implications on normal everyday-life functioning, and more than 50% of suicide attempts occur within a depressive episode [26]. In an attempt to understand and minimize the duration of depressive episodes, researchers within behaviorism have extensively investigated how behavior and depressive symptoms relate [35]. For instance, an extensive systematic review concluded a positive association between sedentary activities (e.g., viewing TV or computer usage) and risk of depression [65]. Oppositely engaging in social environments had a negative association with the risk of depression [61]. More importantly, daily engagement in activities like getting out the door for physical activity, working, social contact, and leisure is a fundamental part of better psychological well-being [33].

These observations between behavior and better health outcome for the individual—embodied as positive activities—motivate the concepts in the therapeutic intervention known as **Behavioral Activation (BA)** [15]. To support BA’s concept, MacPhillamy and Lewinsohn developed the **Pleasant Event Schedule (PES)** [40]. The PES is constructed as a list of 320 pleasurable activities (e.g., “feeding the birds” and “organizing my wardrobe”). Engaging in these activities—or the lack thereof—has been shown to have a significant correlation with depression [36]. Subsequently, the list has played a central role in several interventions, with slightly modified versions to encompass different cultures [3], age groups [51], and disorders [64]. Furthermore, to improve accessibility and better scale BA to the larger population diagnosed with depression [68], the intervention has recently been implemented as smartphone-based apps. A review study identified about 10 smartphone-based systems utilizing BA treatment components [27]. The systems were designed to be used in combination with therapy [38] or as a stand-alone tool [11, 12, 70]. In both cases, the studies demonstrated a reduction in the participants’ depressive symptoms.

Recommending health activities is a core principle in BA, which is typically done by the therapist. However, most BA smartphone applications are passive and do not support a more active recommendation of healthy activities, and they do not adhere to the BA therapy methodology [27]. Moreover, as another core principle in BA, therapists help the patients enact a *diverse* set of activities with difficulty tailored to the patient’s level of depression. However, in most smartphone-based apps, the patient has limited support for activity diversity as they are presented with “raw” Lewinsohn-inspired activities [17, 21, 38]. In its raw form, activities are not organized in categories or difficulty levels appropriate for patients’ current depressive state. As a result, patients have found activities in prior apps too comprehensive or too simple [37].

Previously, we designed a **mobile Health (mHealth)** system for BA, named MUBS, to explore the feasibility of using activity recommendations [53]. MUBS applies content-based knowledge from previous activity registrations to support personalized recommendations. The system includes activities from PES in a unique catalog of 384 pleasurable activities. This catalog organizes activities in six distinct categories and three difficulty levels. Activity diversity was further supported by providing daily inspirations of activities within categories neglected in the past 7 days. The patients received a daily notification commenting on the enacted activity categories and could follow a visualization of their level of accomplishment within each category.

The system’s usability and usefulness have been established in a feasibility study where the patients expressed how the system motivated and rewarded the planning and engagement of more enjoyable activities in their everyday lives [53]. This study examines the clinical effect and in detail investigates the relationship between usage patterns of activity activation and clinical outcome. In summary, the present study contributes the following:

- An analysis of depressive symptoms, as measured by the **Patient Health Questionnaire (PHQ)-8**, from 29 clinically diagnosed depressive patients that used the system for 8 weeks. The results reveal a significant reduction in depressive symptoms.
- A sub-group analysis of patients that displayed a larger improvement in depressive symptoms (*beneficiaries*) and patients with minimal improvement (*non-beneficiaries*). A set of designed **Linear Mixed Effects (LME)** models show several statistically significant variations between *beneficiaries* and *non-beneficiaries* in terms of system usage and activity registrations.
- An entropy measurement that quantifies the distribution of registered activities, which show an association between activity diversity and depressive symptoms.

## 2 BACKGROUND

**Cognitive Behavioral Therapy (CBT)** is a problem-solving therapy methodology that focuses on the ‘here-and-now’ [5]. The following is based on a scenario in which the patient did not go to football practice but instead displayed negative thoughts (e.g., “*the others will be better*”), feelings (e.g., anxious), and behavior (e.g., “*lay in bed*”). The therapist would identify the patient’s overall goal (e.g., “*return to football practice*”) and the associated barriers. The therapist would initiate BA tasks, which include scheduling smaller activities that can assist in the overall goal (e.g., “*buy new football shoes*”). The subsequent cognitive therapy part would then challenge the dysfunctional attitude of the patient. The therapist would carefully sympathize with the patient’s experience (e.g., “*I understand your feelings and concerns that the others might be more skilled than you*”) and then support the patient by identifying and directing faulty cognitive ruminations toward pleasant emotions (e.g., “*having an extra player always benefits a team*”).

From the small scenario, we become acquainted with the vital role of the therapist. The required skills and expertise are connected with the high cost of training and employment in CBT [50]. However, by isolating and comparing the different sub-components of CBT, Jacobsen and colleagues demonstrated that there is no statistically significant difference when practicing the BA therapeutic part alone [28]. BA is centered on activities and is, therefore, a more straightforward therapeutic approach that requires less commitment with the clinic and can be practiced by junior mental health workers with less costly and intensive training [50].

### 2.1 Behavioral Activation

BA aims to increase exposure to healthy behavior’s positive consequence (reinforcement) by increasing reenactment, thereby reduce time dwelling on negative thoughts or feeling depressed [15]. As described by Martell et al. [44] and Lejeuz et al. [34], the standard therapeutic procedure involves three steps across 10–12 one-hour-long sessions. (i) A “data collection” phase, where the patient is told to write down, on an hourly basis, what activity they are engaged in. For each activity, the patient has to evaluate the perceived “Pleasure” (i.e., how enjoyable the activity felt) and “Mastery” (i.e., how much effort it took to accomplish it). These parameters are essential for the patient to identify achievable and straightforward activities that still provide a sense of pleasure [44].

The registrations are usually done for 1 to 2 weeks. (ii) A consultation phase where the patient shares the weekly paper schedules with the therapist. Together, they identify positive reinforcement activities that promote healthy behavior. Such activities are also known as functional activities, in line with the patients’ level of “Pleasure” and “Mastery.” In the same degree of detail, they plan out the next week’s activities. (iii) An assessment phase, where the patient is instructed to follow the plan.

Step (ii) involves a series of tasks to identify overall goals and behavior of value for the patient [34]. In an attempt to reach such behavior, smaller sub-goal activities are identified and scheduled. These activities are ranked in terms of “Mastery” to select activities that the patient can initiate and pursue. Measurable success

criteria (e.g., amount in a week) are established for each activity. Therefore, passively defined activities such as “*visit my sister*” become concrete instructions as “*visit my sister twice a week*.”

The three steps are repeated through re-evaluations with the therapist. For instance, if a weekly sub-goal of “*be in bed before 10pm twice a week*” was achieved, the goal would be increased to 4 days in a week for the next assessment phase.

## 2.2 Technology for Behavioral Activation

Smartphones as a mental mHealth application is a heavily explored area, covering more than 10,000 applications in the digital distribution platforms [67]. The popularity of smartphone-based support in mental health is due to several key advantages: (1) Accessibility to expert knowledge of information and help. (2) Facilitating a reduction in reconstruction bias, as smartphones are within reach 90% of the time with access “in-the-moment” [13]. (3) Improve treatment-seeking behavior as smartphones can limit clinical involvement. This includes reaching patients with stigma toward the clinic or patients preferring anonymity using a mHealth application where reporting symptoms and behavior is perceived as more comfortable [66]. (4) Powerful computational platform with numerous embedded sensors including accelerometer, and GPS location services, which provides information related to human behavior [23]. (5) Computational power enables applications to process data and information from the patient’s subjective inputs or sensors to provide feedback. Feedback can be communicated in many ways and allows for patient empowerment and control during treatment.

An example of a type of empowerment that patients gained was demonstrated in newer research on a BA mHealth application [55]. The application supported the BA treatment methodology from Martell et al. [44]. All the patient’s BA actions in their everyday life was digitalized to run on their smartphones. Patients who had previously been enrolled in a typical BA therapy session used the application in a 4-week feasibility study. The study highlights several advantages in the transformation from paper to a digital application, including notifications, anonymous registrations in public, and the possibility of visual feedback. The visual feedback facilitated a deeper understanding of their depression, regarding awareness of what activities had a positive impact and which to avoid. Deady et al. further investigated visual feedback in their BA application [12]. They developed the application to reduce depressive symptoms through a 30-day challenge program laid out as a step-by-step user journey map. The program consisted of BA tasks, e.g., activity planning, and goal-setting. Completing daily challenges would then be emphasized as a checkmark on the map. A 5-week-long single-arm feasibility study with 84 participants demonstrated a significant reduction in depressive symptoms as measured by PHQ-9. A BA application called Aptivate implemented a badge system, and participants were rewarded when they used a daily calendar to schedule and keep track of activities and daily mood [10]. Aptivate supported the patient to schedule activity values within five life areas (e.g., relationships and health). An example of value would be: “*Call my children twice a week*.” They conducted an 8-week feasibility study comparing usual treatment (i.e., control) with the addition of the application, and demonstrated that participants with access to the application had significantly reduced depressive symptoms over time compared to control. They added on average 21.73 activity values and had a retention rate of 50% after 8 weeks of use. Other BA systems support the patient with specific positive activities as inspirations that are normally coordinated with the therapist [38, 42, 70]. Ly et al. demonstrated that a BA application with support of activity inspirations to increase everyday activities could reduce therapy time by, on average, 47% while maintaining the same treatment outcome [38]. In circumstances that render psychotherapy access impossible, Mantani et al. ran a large **Randomized Control Trial (RCT)** study involving more than 160 patients to investigate the benefit of offering help with an mHealth application as an addition to medicine [42]. They observed a statistically significant reduction in depressive symptoms, corresponding to a 2.48 lower PHQ-9 score after using the application together with medicine for 9 weeks compared to medicine alone. As a unique feature, Wahle et al. [70], supported the recommendation of specific activities that are personalized to the user’s preferences. The app was designed as a stand-alone app without any involvement with the clinic. Every 6 hours,

the users were presented with new activities. Based on a rating system—after they completed the activity—the system re-evaluated what activities to recommend next. A single-arm feasibility study involving 12 participants to use the application for 8 weeks showed a drop in depressive symptoms. The study was interesting, as it was the first to demonstrate a more active role of the mHealth application. The calculated recommendations mimic the therapist’s role to suggest activities that are in line with the patients’ preferences.

Unfortunately, a larger systematic review from 2016 [27] shows that the present applications did not adhere to the core BA techniques. For example, Wahle et al. do support personalized activity suggestions, but lack planning of activities or the goal-oriented concept of BA.

We designed MUBS [53] to support the core BA ingredients. This includes a predefined goal, activity and depressive symptoms tracking, scheduling of single or re-occurring activities, daily reflections, activity inspirations, and content-based recommendations weighted based on context. The application emphasized a personalized design through flexible use-cases and labeled activities according to category and difficulty. The design contribution, the unique personalized recommendation feature, together with an understanding of users’ perception of the technology, has been presented previously [53]. In this study, we provide unique objective insights on the clinical outcome and how usage differed for those patients that displayed an improvement in depressive symptoms. The study utilizes robust second-order longitudinal modeling to capture detailed changes during the study. Such insights are rare [21] as similar works are either (i) technical contributions focusing on the system and the human-computer interaction (e.g., References [4, 20, 57]) or (ii) clinical papers in medical journals that focus on symptom progression (e.g., References [11, 39]).

### 3 STUDY METHODOLOGY

#### 3.1 Participants

Recruitment was done partly via the Psychiatric Center Copenhagen (a hospital clinic) and partly via online channels, including a national patient recruitment site<sup>1</sup> and social media advertisements (e.g., Medium<sup>2</sup>). Recruitment continued beyond the time frame described in the usability study [53] to include 86 participants that downloaded the app. Participants recruited via the Psychiatric Center Copenhagen were clinically diagnosed with depression or bipolar disorder by the clinicians. Participants via online channels self-reported that they had a diagnosis of affective disorder.

Participants from the psychiatric clinic and the national patient recruitment site were invited for a semi-structured interview at the end of the study. They were compensated with an €80/\$100 gift card.

#### 3.2 Procedure

The study was conducted as a one-arm clinical feasibility study, following the recommendations on study design for personal health technology research [31]. Ethical permission to run the study was granted by the Danish National Committee on Health Research Ethics (File no.: H-19002943). Participants who expressed interest in the study were contacted by phone for detailed instructions. They were instructed to use MUBS daily for 8 weeks. Details on how to use MUBS combined with an explanation of BA and why it effectively alleviates depressive symptoms, are available as a psycho-education module in MUBS. On study start-up, patients would go through the module, followed by an informed consent form, a demographic survey, and were asked to fill in an initial PHQ-8 questionnaire on subjectively experienced depressive symptoms in the last 2 weeks. The questionnaire was repeated every second week as a scheduled task visible in the participants “activity overview.” Participants were encouraged to contact the study manager if they experienced any issues.

<sup>1</sup>[http://www.forsogsperson.dk/show\\_ad.php?showit=2654&add=posted](http://www.forsogsperson.dk/show_ad.php?showit=2654&add=posted).

<sup>2</sup><https://medium.com/@cph.cachet/an-app-to-get-inspired-and-activated-start-planning-your-activities-today-8da835d97eac>.

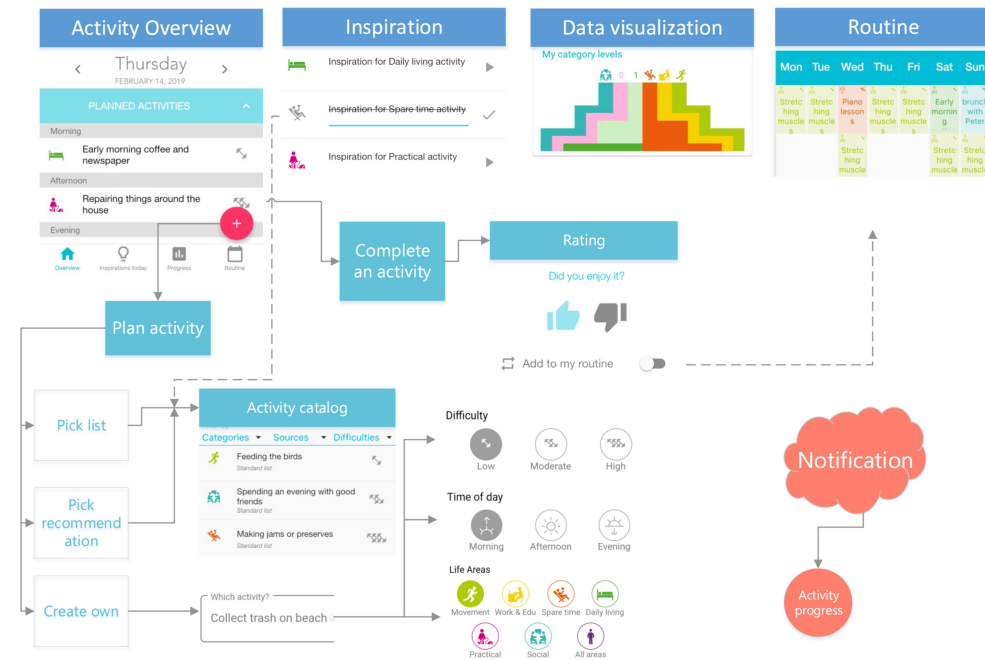


Fig. 1. A flowchart showing some of the main functions in MUBS. According to the design goals, MUBS features a simplified planned *activity overview*, with the possibility to schedule daily or weekly routines. It includes activity inspirations and recommendations through an activity catalog of 384 pleasurable activities. Each planned activity is labeled with an activity category, difficulty, and a time of day divided into morning, afternoon, and evening. Progress within each activity category is shown in “Data visualization” to support activity diversity when planning future activities.

### 3.3 Intervention

Details of MUBS have been published previously [53]. Briefly, MUBS is a smartphone-based personal health technology [4] for BA that provides personalized planning, recommendation, and activation of healthy activities. It was designed in a user-centered design process [25, 43, 60] involving an interdisciplinary group of technical and clinical researchers and a group of patients with depressive disorders. The overall design goal of MUBS is to support the therapeutic methodology of assisting in the planning and scheduling of enjoyable activities, which has been found effective in reducing depressive symptoms. The system is designed to be used daily and work “independently” of the clinic, i.e., it should—without involvement from the therapist—support the planning and enacting of activities by providing personalized recommendations and inspirations from a catalog of enjoyable activities.

A flowchart of the main functionalities in MUBS is shown in Figure 1 and consists of the following:

- **Activity Overview** – MUBS provides a simplified overview divided into morning, afternoon, and evening time slots. The “Pleasure” and “Mastery” scores commonly used for activity rating in BA therapy has been replaced with a binary “thumbs up or down” and a difficulty between 1 and 3 dumbbells, respectively. The difficulty became an “expected” value prior to engaging in the activity, which has previously been demonstrated better to reflect users’ depressive symptoms [22]. This difficulty level then enables the patient to sort and filter activities.
- **Activity Catalog** – The patients can access a catalog of 384 enjoyable activities when planning an activity. The catalog is an extension of the 320 activities from the PES work of Lewinsohn and MacPhillamy [41]

with novel activities presented by Mørch and Rosenberg [46]. Each activity is labeled with the difficulty level and categorized as “movement” (physical activities), “work & education” (activities involving learning or work), “spare time” (i.e., leisure activities, usually hobbies or other activities that are done alone), “daily living” (i.e., basic self-care activities such as “Eating” and “Showering”), “practical” (Practical activities covering “cleaning” and “buying groceries”), or “social” (i.e., activities with others, which also covers virtual settings such as a phone call, or video chat). When the participants created their custom activity, they had the possibility to select a seventh activity category, “other,” in case none of the previous categories were representative.

- **Activity Rating** – When the patient had done an activity, s/he would rate it with a “Pleasure” score. A Naive Bayes recommender model [54] utilizes these “Pleasure” scores to identify a set of activities from the catalog that the user most likely would enjoy. Finally, when activity was rated, the patient had the option to add this activity as a weekly routine activity.
- **Data Visualization** – Two types of visualizations guide the user toward more balanced and diverse weekly activation. One type displays the total amount of registered activities distributed within each category. The other, as shown in Figure 1 on the “Data Visualization” screen, displays a three-step staircase picture for each activity category, aggregated together like a pyramid. The patient’s placement in the pyramid was determined by three unique threshold values derived from a linear combination of the number of activities done in the past 7 days and their respective difficulty level weighted by a category dependent constant. The constant was derived from detailed hourly activity registrations of patients labeled activities in the self-same categories [55].

### 3.4 Data Processing

All data sampled from the app is stored in Google Firebase. MUBS applies record level encryption with AES-256 and HTTPS/TLS for all patient data transmissions and provides secure indexing and tokenization for each device. The encryption keys are managed according to HSM standards. A custom Javascript script is used to extract all registered activity and PHQ-8 data into local JSON files for further data analysis. We calculated summary statistics describing the registered activities in Python (v. 3.7). We then used Matlab (v. R2018b) to model the activities and PHQ-8 levels across the study. A LME model was developed to capture the co-variance within each patient. Control for individual variation among the patients was done by adding random effects ( $B_0, B_1$ ).

Specifically, for PHQ-8, we were interested in knowing whether the fixed effect of time ( $\beta_1$ ) interacted with PHQ-8 and whether there was any quadratic interaction ( $\beta_2$ ). The full model, which includes the quadratic interactions, for PHQ-8 can be written as follows:

$$PHQ(t) = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot t^2 + B_0 + B_1 \cdot t + \epsilon(t), \quad (1)$$

where  $\epsilon(t)$  represents the error function.

A combination of a simulated Likelihood ratio test and inspection of the conditional residuals between the full and reduced fitted models, e.g., models without the quadratic term and random effects, was used to determine the included factors in the final model specification. We used the individual variation across the study period ( $B_1$ ) to separate patients into two distinct sub-groups. The first sub-group comprises of patients who experienced a larger drop in PHQ-8 compared to the general representation of the study population ( $B_1 \leq 0$ ), denoted as *beneficiaries*, and the second sub-group that displayed less-to-none improvement of depressive symptoms ( $B_1 > 0$ ), which we represent as *non-beneficiaries*. A second PHQ-8 model was designed to investigate the study trajectory for patients with moderate to severe depressive symptoms and patients with mild depressive symptoms at study onset. Here, a binary “dummy” vector (1, if the patient had moderate to severe symptoms at study onset, 0 otherwise) together with its interaction with time and the quadratic time term was added to the model in Equation (1).

Continuing with the separation of *beneficiaries* and *non-beneficiaries*, we defined LME models to analyze the difference between the two groups concerning app usage, including the number of registered activities, the proportion of planned activities and the number of activities picked from the activity catalog, and the composition and change in activity difficulty and activity categories. Denoting these dependent variables as  $Y(t)$ , the model can be written as follows:

$$Y(t) = \beta_0 + \beta_3 \cdot Grp + \beta_4 \cdot Age + \beta_5 \cdot Sex + \beta_6 \cdot A_T + \beta_1 \cdot t + \beta_7 \cdot Grp \cdot t + \beta_2 \cdot t^2 + B_0 + B_1 \cdot t + \epsilon(t). \quad (2)$$

$Grp$  and  $Sex$  are binary variables defining those that are *beneficiaries* and male, respectively, and  $A_T$  is the total amount of activities that the specific participant has registered during the study. We centralized time,  $Age$ ,  $Sex$ , and  $A_T$ , e.g., subtracting the average value from the variables, to ease the interpretability of the weightings.  $\beta_4$  to  $\beta_6$  have been included in the model to adjust for the covariates of  $Age$ ,  $Sex$ , and  $A_T$ . Unadjusted models (i.e., without  $\beta_4$  to  $\beta_6$ s, and random terms) were also analyzed.

The change in activity categories was merged into a mutual information entropy calculation. Based on Shannon's entropy, the calculation measures how uniform ("even") the distribution of the activities categories was for the registered activities. The entropy of the activity categories,  $H(AC)$ , was defined as follows:

$$H(AC) = - \sum_{i=1}^M F(x_i) \log_e F(x_i) \quad F(x_i) = \frac{1}{N} \sum_{j=1}^P AC_{x_i}(j). \quad (3)$$

$F(x_i)$  is the occurrence frequency for a specific activity category  $x_i$ .  $M$  is the total number of activity categories,  $N$  is the total amount of activities, and  $P$  is the number of participants. Therefore,  $AC_{x_i}(j)$  represents the total amount of activities for a given activity category and participant.

## 4 RESULTS

Many participants who downloaded the app did not use the app for more than 1 week ( $57/86 = 66\%$ ). From these, 75% never registered any activities ( $43/57$ ). The dropouts had a mean age of 32.7, the range was 16–57 years, and 63% were female. Table 1 presents an overview and analysis of the dropouts and the patients included in the subsequent analysis. Demographic and initial depressive symptoms did not differ between the dropouts and the remaining sample of participants as appropriately assessed by either the z-score from a non-parametric Mann–Whitney  $U$  test or the chi-square test. However, there was a significant difference in the recruitment channel ( $z = 36.66, p < .001$ ). A majority of the dropouts included participants who found study details on MUBS through social media. We report results from patients who used the system daily, resulting in 29 eligible patients with sufficient data for the subsequent analyses.

### 4.1 Clinical Outcomes

The bi-weekly Patient Health Questionnaire (PHQ)-8 scores were used to model changes in depressive symptoms. Figure 2(a) shows the overall fitted model (bold line with shadowed 95% Confidence Interval (CI)). Through the Likelihood Ratio test, the quadratic model was deemed as a better fit. This was also verified by inspecting the residuals (RMSE = 1.89) and directly comparing the fitted values with the measured (Figure 2(b)). We mapped the "Fitted" model's values against the observed "Response" PHQ-8 values. We see that the points are evenly split around the diagonal slope with no apparent trend or pattern.

The linear and quadratic fixed effect of time yielded an F ratio of  $F(1, 127) = -3.67, p < .001$  (CI  $[-2.22, -0.66]$ ), and  $F(1, 127) = 2.90, p < .005$  (CI  $[0.03, 0.19]$ ), respectively. The fitted model indicates that the patients experienced a statistically significant reduction in PHQ-8 score ( $\beta_1 = -1.44, SD = 0.39$ ) every week, but the rate of change was slowed by the quadratic term ( $\beta_2 = 0.11, SD = 0.04$ ). As we included the random effect of time, we could depict which patients experienced an improvement in depressive symptoms during the study period.



Table 1. An Overview and Statistical Comparisons between the Included Sample and the Dropouts

	Sample		Dropouts		Statistics		<i>p</i> value
	Mean or total	SD or <i>n</i> (%)	Mean or total	SD or <i>n</i> (%)	$\chi^2$	<i>z</i> -score	
N (% Female)	29	72%	57	63%	0.49		.58
Age (years)	33.38	10.47	32.67	10.12		-0.23	.82
Social media recruit	10	34%	54	95%	36.66		<.001
PHQ-8 study onset	10.83	6.51	13.13	6.47		1.59	.11
<i>Occupation</i>							
Full time	8	28%	26	46%	2.61		.11
Student	9	31%	16	31%	0.08		.77
Unemployed	9	31%	10	18%	2.03		.15
Part time	3	10%	3	5%	0.77		.38
Leave of absence	0	0%	2	4%	1.04		.31

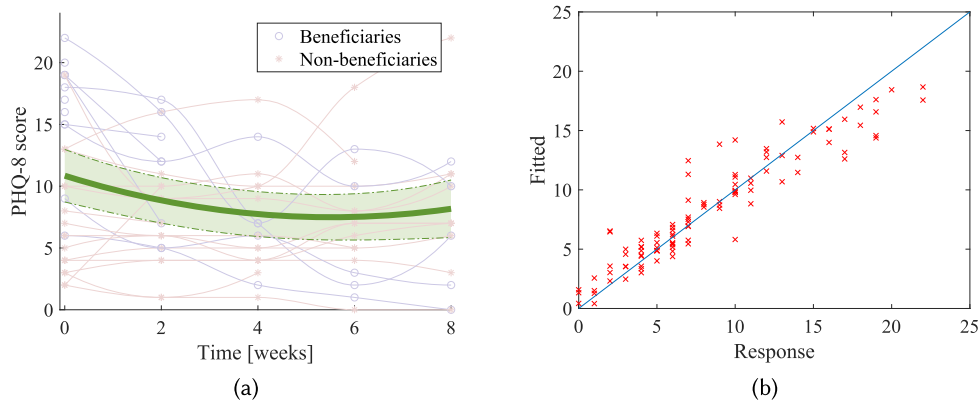


Fig. 2. PHQ-8 as a function of time. (a) The bold green line is the fitted LME model for the PHQ-8 responses as a function of study duration. The area within the dotted lines represents the 95% CI. The observed PHQ-8 values across time for each patient is included as a Spaghetti plot. Patients with a random slope effect lower or equal to zero are highlighted in blue and defined as *beneficiaries*; patients with a positive random slope effect are highlighted in red and defined as *non-beneficiaries*. (b) The fitted values from the model against the observed PHQ-8 responses.

The patients who had a random-effects slope value below or equal to zero ( $n = 12$ ) were defined as *beneficiaries*, while the remainder ( $n = 17$ ) were defined as *non-beneficiaries*. These two distinct groups of patients are shown as Spaghetti plot lines of their observed PHQ-8 scores in Figure 2(a) in blue and red, respectively.

A large portion (83%) of the *beneficiaries* were patients who experienced moderate to severe depressive symptoms ( $\text{PHQ-8} \geq 10$ ,  $n = 15$ ) at study onset. Therefore, we expanded the previous model to include a nominal identifier that separates patients based on their symptoms at study onset. The other patients reported minimal to mild depressive symptoms ( $\text{PHQ-8} < 10$ ,  $n = 14$ ) at study onset. The fitted model for both groups is shown in Figure 3(a). The blue lines are patients with moderate to severe depressive symptoms, while red is the patients with minimal to mild depressive symptoms at study onset. The bold line is the fitted model for the group, with the corresponding 95% CI in dotted lines. The observed PHQ-8 values for each patient are represented as faded lines. The residuals (RMSE = 1.92) as fitted values versus the observed responses are shown in Figure 3(b). While there was significant overlap in *beneficiaries* and patients who experience more moderate symptoms at enrollment, the latter includes patients with unchanged conditions or worsening depressive symptoms during the

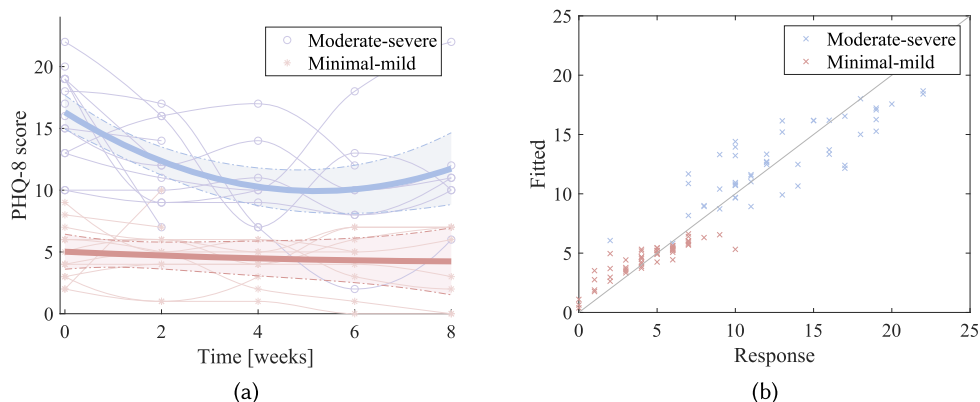


Fig. 3. PHQ-8 as a function of time separated in patients with moderate to severe and minimal to mild depression at study onset, respectively. (a) The bold blue and red lines are the fitted LME model for the PHQ-8 responses as a function of study duration for the two patient groups. The lighter shaded area indicates the 95% CI. The observed PHQ-8 values across time for each patient is included. (b) The fitted values from the model against the observed PHQ-8 responses. We color-coded the values in blue and red to distinguish participants that initially exhibited moderate to severe depressive symptoms from participants with minimal to mild depressive symptoms, respectively.

study. The remainder of this section will provide a more detailed analysis of the differences between *beneficiaries* and *non-beneficiaries*.

## 4.2 Group Differences

Overall statistical differences between the *beneficiaries* and *non-beneficiaries* are summarized in Table 2. The table is separated into three domains, (i) demographics, (ii) app usage, and (iii) reported activity differences. There were no statistically significant differences in patient demographics. Sex and occupation (full-time), as categorical data, were assessed with the chi-square test, while we used the Mann-Whitney  $U$  test for the age differences. We applied two different LME models, an unadjusted version with no random effects, equivalent to a General Linear Model, and an adjusted model. The adjusted model includes random effects to account for individual differences in the dependent variables and controlling for age, sex, and the total number of activities registered. In the subsequent sections, we will report results from the adjusted model unless otherwise specified.

## 4.3 App Usage

The number of activities was very similar across the two groups ( $\beta_0 = 0.33, p = .91$ ), although it seems to drop more during the study for the *beneficiaries* ( $\beta_1 = -3.93, p = .10$ ), albeit not statistically significant. Similarly, the number of planned activities beforehand dropped more for the *beneficiaries* during the study ( $\beta_1 = -4.56, p = .07$ ).

In the app, the patients could get inspired from a catalog of pleasurable activities based on Lewinsohn's PES. On average, the *beneficiaries* used the catalog to a lesser extent ( $n = 57.25$ ) than the *non-beneficiaries* ( $n = 73.59$ ). The difference was statistically non-significant ( $\beta_0 = -6.29, p = .09$ ). However, as the study progressed, the *beneficiaries* selected fewer activities from the catalog than the *non-beneficiaries* ( $\beta_1 = -3.92, p = .03$ ). Figure 4(a) shows the statistically significant change, and we can observe that both sub-groups initially select ca. 20 activities from the catalog, but throughout the study, the *beneficiaries* choose fewer catalog activities, while *non-beneficiaries* were stable. Since both sub-groups had the same number of activities (cf. Table 2), the *beneficiaries* added statistically significantly more personal, custom-made activities ( $n = 56.25, \beta_0 = 7.86, p = .02$ ), instead. While there was

Table 2. Summary Statistics for the *beneficiaries* (B) and *non-beneficiaries* (nB) Patients

Description	Value		Unadjusted Model		Adjusted Model <sup>‡</sup>	
	B	nB	baseline, $\beta_0$	interaction, $\beta_1$	baseline, $\beta_0$	interaction, $\beta_1$
N (% F)	12(75%)	17(71%)	$\chi^2 = 0.07$	—	—	—
Age	33.67	33.18	$z = 0.29$	—	—	—
Full time	5 (42%)	3 (18%)	$\chi^2 = 2.03$	—	—	—
Activities	113.50	110.41	5.14 [−7.6:17.9]	−1.35 [−12.4:9.7]	0.33 [−5.8:6.4]	−3.93 <sup>†</sup> [−9.3:1.4]
Planned	86.67	69.29	8.14 <sup>†</sup> [−2.6:18.8]	−2.25 [−11.6:7.1]	4.49 [−3.4:12.4]	−4.56 <sup>†</sup> [−9.5:0.4]
Catalog	57.25	73.59	−3.20 [−12.1:5.7]	−2.34 [−10.1:5.4]	−6.29 <sup>†</sup> [−13.6:1.0]	−3.92 <sup>*</sup> [−7.4:−0.4]
Custom	56.25	36.82	8.34 <sup>*</sup> [1.3:15.4]	0.99 [−5.1:7.1]	7.86 <sup>*</sup> [1.2:14.6]	1.21 [−4.7:7.1]
Routine	62.33	42.18	8.03 [−2.5:18.5]	−3.17 [−12.3:6.0]	4.64 [−5.6:14.8]	−6.11 <sup>†</sup> [−12.3:0.1]
Difficulty 1	49.67	60.12	−1.25 [−11.3:8.8]	−0.52 [−9.3:8.2]	−3.50 [−10.9:3.9]	−1.51 [−4.5:1.5]
Difficulty 2	35.58	30.59	2.65 [−1.9:7.2]	−1.29 [−5.3:2.7]	0.75 [−4.9:6.4]	−2.42 <sup>†</sup> [−5.4:0.5]
Difficulty 3	28.25	19.71	3.75 <sup>*</sup> [0.5:7.1]	0.46 [−2.4:3.3]	1.68 [−3.4:6.8]	−0.90 [−2.5:0.7]
Movement	7.67	17.76	−2.77 <sup>**</sup> [−5:−0.5]	−0.12 [−2.1:1.8]	−2.75 <sup>†</sup> [−5.8:0.3]	0.09 [−1.2:1.3]
Work	27.92	14.29	5.30 <sup>**</sup> [2.1:8.5]	0.61 [−2.2:3.4]	5.54 <sup>†</sup> [−0.3:11.3]	−1.24 <sup>†</sup> [−2.8:0.3]
Spare time	18.25	30.82	−3.32 [−8.9:2.3]	−1.53 [−6.4:3.3]	−4.56 <sup>†</sup> [−9.5:0.3]	−1.94 <sup>*</sup> [−3.8:−0.1]
Daily living	27.08	14.71	4.6 <sup>†</sup> [−0.8:10.0]	−0.92 [−5.6:3.8]	3.53 [−2.6:9.7]	−1.61 <sup>†</sup> [−3.8:0.5]
Practical	12.67	12.59	.56 [−1.2:2.3]	0.28 [−1.2:1.8]	0.24 [−1.5:2.0]	0.08 [−1.1:1.3]
Social	15.75	19.12	−1.76 [−5.7:2.2]	−0.48 [−2.3:1.3]	−1.72 [−5.1:1.6]	−0.37 [−1.8:1.1]

<sup>‡</sup>Adjusted for sex, age at study onset, total registered activities, and accounting for variability between patients (random effects).

<sup>†</sup>.05 ≤  $p$  ≤ 0.1 \* $p$  < .05 \*\* $p$  < .01.

“Value” is the average number for each patient across the study period. Each value was subject to a LME model, modeled with the same time interval as the PHQ-8 scores (every second week). Baseline and interaction fixed effects correspond to the difference between *beneficiaries* and *non-beneficiaries* halfway through the study (weeks 4 and 5) and the rate of change with time, respectively.

a large difference in the number of custom made activities, we were not able to verify a statistically significant growth in the variation during the study period ( $\beta_1 = 1.21, p = .68$ ).

The patients also had the in-app option to schedule activities as recurring routine activities. Like the number of planned activities, the *beneficiaries* had a larger non-significant amount of routine activities ( $\beta_0 = 4.64, p = .37$ ). As the activities registered dropped compared to the *non-beneficiaries*, the number of routine activities also decreased borderline significantly during the study period ( $\beta_1 = -6.11, p = 0.05$ ). The following section investigates the differences in enacted (or neglected) activities between *beneficiaries* and *non-beneficiaries*.

#### 4.4 Activity Differences

As explained in Section 3.3, the extended Lewinsohn’s activity catalog was labeled with two features: (i) A “difficulty rating” from 1 to 3 that mimics the “mastery” score of BA. A score of 1 indicates an activity that requires minimal effort to engage in, while an activity with a score of 3 requires a more considerable effort to initiate and carry out. (ii) An “activity category” putting each activity into one of six distinct categories. Both features could be customized by the user for each activity, thereby making an activity more personal. However, we only recorded few cases (2.4%) where the patients changed the catalog’s default values.

Looking at the difficulty level of activities, we did not observe any statistically significant difference across the two sub-groups. Although, as shown in Table 2, the *beneficiaries* had more “difficulty 3” activities in the unadjusted model ( $\beta_0 = 3.75, p = .03$ ). This observation, however, was not significant when adjusting for age, sex, total activities, and controlling the variability between the patients different onset values ( $\beta_0 = 1.68, p = .52$ ). The change was not influenced by the age ( $\beta_4 = -0.05, p = .73$ ) and sex ( $\beta_5 = 0.11, p = .97$ ) covariates, but

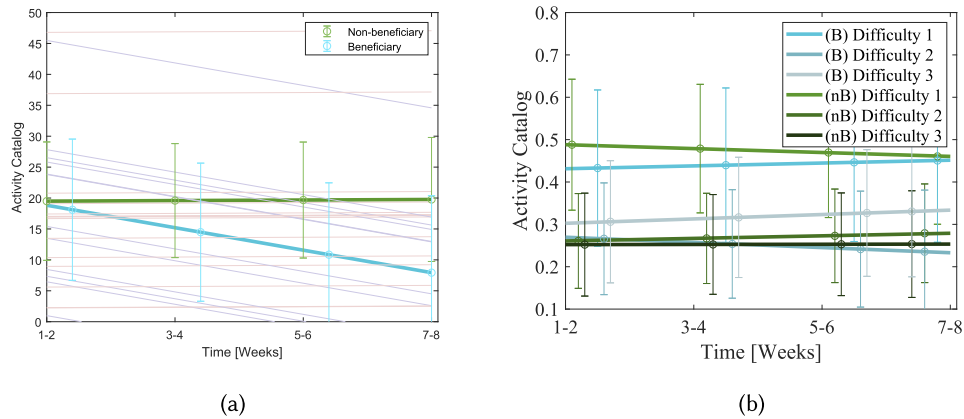


Fig. 4. (a) The adjusted LME model for the activities picked from the Lewinsohn catalog of enjoyable activities. The model for *beneficiaries* is visualized in blue, with each patient in a faded blue shade, The *non-beneficiaries*' model is shown as green regression line with each patient as a faded shade of green. The 95% CI is shown for four time-points as error bars. (b) The LME model for the different difficulty levels separated in *beneficiaries*, illustrated in shades of blue, and *non-beneficiaries* in shades of green.

a result of the variation among the patients ( $B_0 = 6.45$ , CI [4.8, 8.7]). Figure 4(b) illustrates the changes in difficulty during the study. The numbers have been normalized with the total number of activities such that the sum of difficulties equals 1. In alignment with Table 2, there is a larger average amount of “difficulty 1” activities in the *non-beneficiaries* sub-group and a more considerable amount of “difficulty 3” in the *beneficiaries* sub-group.

Table 2 also shows the distribution of activities across the activity categories. In raw numbers, the *beneficiaries* had more “work & education” ( $n = 27.92$ ) and “daily living” ( $n = 27.08$ ) activities. Particularly in the unadjusted model, there is a statistically significant higher amount of “work & education”-related activities in the *beneficiaries* group ( $\beta_0 = 5.30$ ,  $p < .001$ ). This difference was borderline significant in the adjusted model ( $\beta_0 = 5.54$ ,  $p = .06$ ). A similar pattern is observed for “movement” activities. Here, in the unadjusted model, the *non-beneficiaries* had statistically significant more activities in this category ( $\beta_0 = -2.77$ ,  $p = .01$ ) but the significant effect did not carry over in the adjusted model ( $\beta_0 = -2.75$ ,  $p = .08$ ). There is no difference in the number of “social” activities between the two sub-groups ( $\beta_0 = -1.72$ ,  $p = .31$ ) or in how this amount changed during the study ( $\beta_1 = -0.37$ ,  $p = .58$ ). However, there is a statistically significant increase in “spare time” activities for the *non-beneficiaries* during the study ( $\beta_1 = -1.94$ ,  $p = .04$ ).

Figure 5 shows the activity categories’ distribution to visualize the diversity of activities between the two groups. The bar charts are grouped for each PHQ-8 2-week assessment period. In line with Table 2, we observe that “spare time” activities increase during the study period for the *non-beneficiaries* while remains stable for the *beneficiaries* (third bar for each period). The “work & education” bar (second bar) is higher in the *beneficiaries* for all time points, mostly driven by the large number of activities registered as “Working” or “Writing diary entries” from the activity catalog. Figure 5 also shows the entropy calculations for each period. Interestingly, we observe a large drop in entropy for the *non-beneficiaries* during the study. The decline is mostly driven by the chaotic shift toward a larger quantity of “spare time” activities, with registrations such as “Watching TV or videos” and reduction of “work & education”-related activities. Oppositely, the *beneficiaries* display an increase in entropy during the study. The shift includes a borderline significant lowering of “daily living.” The smaller “daily living” drop is caused by a reduction of more mundane type of activities such as “Putting moisturizing cream on my face/body” and “Waking up early or/and getting ready at a leisurely pace.”

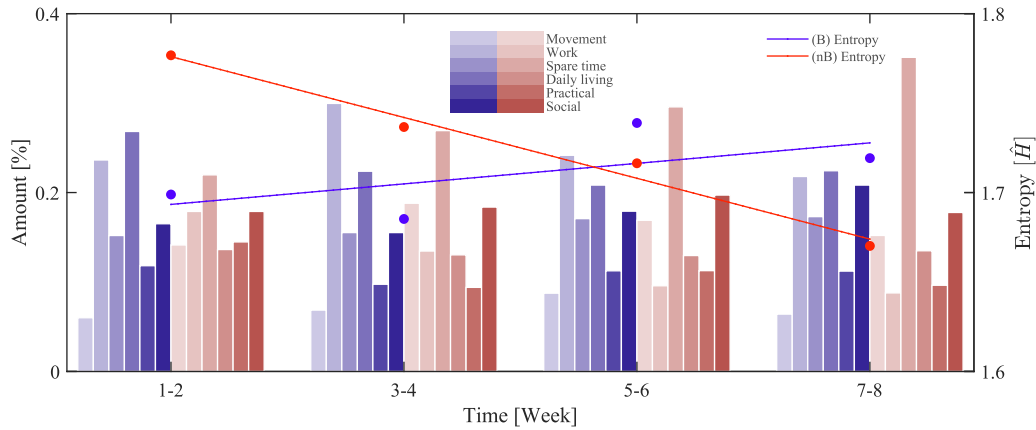


Fig. 5. A bar chart distribution of the six activity types used by the patients during the 8-week study period, divided into the *beneficiaries* (blue) and *non-beneficiaries* (red) patient sub-groups. The overall entropy of each 2 weeks is marked with a circular dot for each sub-group. A least-square fit of each sub-group’s entropy is shown as the dotted line in the respective colors.

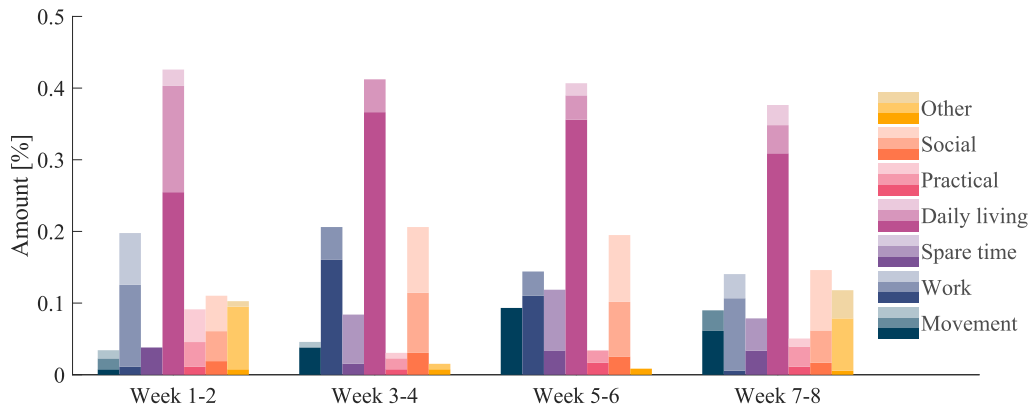


Fig. 6. Distribution of the *beneficiaries* custom created activities within the different activity categories. Further, the activity category is split in difficulty from 1 to 3, respectively (bottom, the solid color part is less difficult).

The adjustments observed were partly exercised by incorporating custom activities. Figure 6 provides an overview of the activity types that the *beneficiaries* created during the 8 weeks. Each bar represents an activity category and is divided according to the amount (in %) of activities within each difficulty level. Several observations can be made from this graph. First, we observe that the distribution of personalized activities is most stable for “work & education” and “daily living.” Second, “daily living” activities are the most common type of activity that is personalized. Third, of the “daily living” activities, most activities are labeled as “difficulty 1,” as shown by the darker color scheme. Examples of custom activities include “*Take all scheduled meds,*” “*Prepare tomorrow’s meals,*” “*Go to bed by 10pm,*” “*Pack gym bag,*” and “*Eat a healthy breakfast.*” Additionally, we could register that a large majority of “daily living” activities were routine (e.g., scheduled as recurring) activities ( $257/281 = 92\%$ ) followed by “work” ( $74/121 = 61\%$ ), which stand in contrast to the others (movement: 14%, spare time: 10%, practical: 34%, social: 19%, and other: 2%).

## 5 DISCUSSION

Although BA is a lightweight and simple treatment methodology for depressive symptoms, it has repeatedly been proven to be a promising treatment option compared to other CBT methods [9]. Due to its simple techniques, BA has also been supported in smartphone-based applications with positive outcomes (e.g., References [10, 12, 21]). This study similarly showed a statistically significant reduction in depressive symptoms across the 8-week use of the system. Moreover, this study further provided detailed insights into how usage and registered behavior differed for the sub-group of patients, displaying a statistically significant reduction in depressive symptoms.

Notably, many enrolled participants downloaded but never used the system (50%), resulting in a much lower factual dropout ( $14/43 = 33\%$ ). We demonstrated that the dropouts did not differ regarding initial depressive symptoms or other demographics. Instead, a significant amount of the participants found the system through social media, therefore not included for the end interview and following financial compensation. Given this, the numbers observed should be treated as a “real-world” situation. Fleming et al. made a large systematic review of “real-world” uptake of self-help interventions for depression [19]. Surprisingly, they observed that users using such systems at least once after downloading range between 21% to 88%. As an example, only 61.1% returned to use a PTSD coach app after installing it [19]. Dahne et al. compared their study population between remotely or locally recruited. They found a significantly lower number of app sessions for the remotely recruited participants [10]. Both studies did not discuss why these numbers were low, and we can only speculate on whether it is due to, e.g., an initial curiosity or something related to the system design. However, the strength is our sample of diagnosed patients in every possible symptomatic depression state where most studies filter their study population to only include participants with more moderate depressive symptoms at study onset (e.g., References [6, 11, 70]).

In our heterogeneous sample, we found several common traits that differentiated patients that improved the most within the study period, which we defined as *beneficiaries*. These patients reduced their use of pre-made activities from the catalog and seemed to develop a more *personalized* use, including creating their own custom activities. Furthermore, they appeared to display a more *diverse* activation using activities more evenly distributed among the different activity categories.

While discussing the results in more detail, we would like to emphasize that the results are drawn from a one-armed feasibility study. Without a RCT, comparing with a control group not having access to MUBS, this study does not allow for causal inference to verify whether the *beneficiaries* and their usage adaptations and behavior patterns can, in fact, be attributed to the use of MUBS. Therefore, the result and insightful associations between *beneficiaries* and their particular usage of this type of mHealth technology should be used to guide future studies concerning the efficacious design and adaptation of therapeutic interventions.

### 5.1 Clinical Outcome

Participants showed a statistically significant reduction in PHQ-8 scores while using MUBS (Figure 2(a)). The decline in depressive symptoms was even more apparent when we analyzed a sub-group with moderate to severe depressive symptoms at the study onset, as shown in Figure 3(a). Similar observations, including the quadratic effect, have been reported in other studies utilizing a mobile intervention to target depressive symptoms [42, 70]. The more considerable decline for these participants indicates that the system has a more pronounced effect for users experiencing more severe depressive symptoms. Such an outcome would be in line with findings from traditional BA therapy. In traditional therapy, it is established that BA significantly outperforms other cognitive therapy methods among more severely depressed patients [15]. The treatment advantage for more complex, severe depression has not been established in mHealth applications. Yet, mHealth studies with participants experiencing moderate to severe depressive symptoms at baseline (e.g., Reference [10]) demonstrate a significant treatment effect. However, these observations of reducing depressive symptoms in our and other studies may simply be regression toward the mean. Moreover, as mentioned previously, the study lacks a control group (a second arm

that did not have MUBS installed). Even though this study points to the feasibility of using activity recommendation technology for depressive patients, we cannot conclude that the effect can be assigned to the use of MUBS alone, since other factors may have contributed to the reduction in depressive symptoms for the *beneficiaries*.

Nevertheless, and more interestingly, the study provided detailed quantitative insight into the usage of MUBS and the kind of activities registered, which can help understand—from a behavioral perspective—how this kind of mHealth technology for mental health can benefit and help depressed patients. The main findings here was that activity personalization and diversification seemed to benefit patients.

## 5.2 Personalizing Behavioral Activation

Lewinsohn and his work on the PES have been a central aspect of BA therapy [29]. The method to re-directing the patients' behavior into activities that elicit positive reinforcement is a crucial path toward improvement in depressive symptoms [44]. These BA techniques have been incorporated into mHealth systems as recommendations [70], assistance [42], or inspiration [38]. A qualitative study of MUBS similarly discovered that patients found inspiration from specific activities, and in particular, planned activities seemed to facilitate activation [53]. Hence, all prior studies found that the PES or a similar activity catalog is central for BA mHealth systems.

However, this study showed that the activity catalog was used less as the *beneficiaries* progressed with the app and reported fewer depressive symptoms during the study period (see Figure 4(a)). The *beneficiaries* reduced uptake from the activity catalog can be attributed to the increased personalization of MUBS. MUBS has been designed to adapt and allow for personalization, including the support for (i) adding and modifying activities, (ii) sorting activities in categories and effort, (iii) content-based recommendations as an addition to browsing the activity catalog [53], and (iv) creating recurring “schedules” of activities. The ability to create recurring events as daily routine patterns support standard behavior and have several therapeutic advantages to the benefit of multiple patient groups [62]. Routine patterns, also represented as circadian rhythms, are traceable to the fundamental body- and neuro-behavioral processes, including core temperature, systolic blood pressure, cortisol, urine volume, sleep, and mood [1, 32]. The effect on mood is so prominent that circadian dysregulation is believed to be central in depressive and manic episodes [24]. Recently supported in a study on activity logging from patients undergoing BA therapy, patients who followed their circadian activity rhythm reported a higher score of daily “pleasure” [56].

While we did not investigate circadian activity rhythms directly, we instead observed how personalization regarding two types of custom created activities scheduled was constant throughout the study period. This is illustrated in Figure 6, where the distribution of personally engaged “daily living” and “work”-related activity types remained unchanged. The lack of change could be ascribed to the activities being scheduled within a routine framework and contributed to the *beneficiaries*' positive outcomes. We observe that the type of activity most personalized within such routine behavior was “daily living.” This is in line with Nunes [48], who argues that self-care technologies should focus on supporting more mundane everyday tasks, such as remembering to take medication. In this study, we observed how several *beneficiaries* patients created activities such as “*Take medication*” or “*Take all scheduled meds*” and how these activities were interchanged among morning, afternoon, and evening. Furthermore, a simple mundane “daily living” activity such as “*Brush my teeth*” was added by *beneficiaries*. Hence, from this use of MUBS, we see the impact of customizing and planning simple activities. Mundane activities may serve as essential building blocks in maintaining routines in everyday life and burdensome days, serving as valuable positive reinforcement.

Personalization has also been supported in other mHealth systems [54]. Rabbi et al. [49] provided personalized activity suggestions to promote physical activity, positively received by the test users. Wahle et al. [70] separated activities into categories and provided recommendations within these categories based on past ratings. Personalization was slowly integrated as less relevant activity categories were rated low and phased out. Similarly to MUBS, Springer et al. [63] included the labeling of activities in categories. As the system learned what

categories positively impacted mood, the system suggested activity categories to schedule that would change mood accordingly. The system did not include specific activities but demonstrated the advantage of a personal model contrary to a population-based model. While these other systems do not include BA activity scheduling as the personalizing feature, they coincide with this study’s finding that personalization has significant implications for the difference between *beneficiaries* and *non-beneficiaries*.

### 5.3 Activity Diversity

It is known that a daily display of activity diversity involving a good night’s sleep, remember medication, refresh oneself with a shower, get out the door for a walk, and engage in some social activity is important for psychological well-being [47]. However, only a few studies have directly investigated how activity diversity affects depressive symptoms and mood [30, 33]. Similarly, studies have shown that excessive, i.e., non-diverse, behavior can affect mental well-being. Dhir et al. [14] show that excessive media use can harm a person’s well-being and mental health. Saeb et al. [59] have shown how a reduction in location entropy, i.e., a reduction in the number of visited places, is strongly correlated with depressive symptoms, and Rost et al. [58] show that excessive mobile phone usage harms mental well-being.

We used Shannon’s entropy to measure activity diversity during the study to investigate whether diversity is exercised and reported in our sample. As shown in Figure 5, we observe that the *beneficiaries* achieve a larger entropy, i.e., a more even distribution of activities during the study. The *non-beneficiaries* displayed an opposite change resulting in a lower entropy, i.e., a less diverse selection of activities. On the contrary, the activity distribution is less distributed when we separate and investigate the custom activities alone (Figure 6). From previous qualitative work on MUBS, we learned that custom activities are either created, because the patient has to do something particular, personalized to their surroundings, or that they received inspiration from the recommendations or catalog to make a similar but customized variant [53]. We believe the latter case has been a motivation for incorporating more neglected activity areas (i.e., areas in the lower step of the app’s pyramid visualization). Primarily the areas “movement” and “social,” which ensures a more even distribution of activation when added with the picked activities (Figure 5). Combining the previous studies with our findings—that a more diverse activation seems to benefit mental health—we suggest future studies to be attentive to the connection between exercising a diverse activity level and the resulting positive impact on depressive symptoms.

When observing the average amount of registrations (Table 2), we saw that the *beneficiaries* had an increased amount of “work & education,” and “daily living” activities. The *non-beneficiaries* had more “movement,” and “spare time,” while “practical” and “social” activities were comparable in the two groups. Regarding how these differences were interacting over time (Figure 5), only “spare time” activity had a statistically significant association favoring an increase over time for the *non-beneficiaries*. Some of these observations are in line with the literature. Maintaining a normal work routine, including full-time educational activities, is beneficial for mental health in general [8]. Ensuring time for “daily living” activities indicates that the patient could devote time for self-care, like cooking meals, general hygiene, and sleep [44]. However, we found no significant difference in “movement” and “social”-related activities, and the observed difference favors the *non-beneficiaries*. This seems at odd with prior studies, since increased physical and social activity has repeatedly been found to be a vital marker for the reduction in depressive symptoms (see References [7, 52, 61] for a review). Although most findings are drawn by objective markers from smartphones and wearables, the relation is still represented in subjective registrations [45, 61], as in our study setup. One possible explanation resides in the granularity of details present in the behavioral logging within MUBS. Oppositely to hourly activity sampling, where all daily chores and actions are inevitably reported [56], the patients are presented with three time periods. Consequently, everyday activities such as work might have less personal value to register. By contrast, recommended or unusual activities would make more sense for the user to remember and register. For MUBS to represent behavior in line with the literature, equivalent to hourly activity registrations, a future step could include location-based context sampling. Location context can then be used to derive semantic labels of current activities automatically [69]. However,



it is important to keep in mind that a core principle of BA's therapeutic approach is the active involvement of patients in planning and evaluating personalized activities. Therefore, automatic activity labeling should only be considered as a supplement. In return, such detailed information may improve the system's recommender functionality but simultaneously challenge ethical practices for these future personalized mHealth systems.

As a final note, in the case of the version of MUBS employed during this study, all machine learning operations were conducted in the cloud—a choice made to facilitate ongoing analysis of the technology's functioning and more broadly facilitate the conduct of this research study. In doing so, we took steps to prioritize the privacy and security of users' data. In looking to the secure future up-scaling of MUBS and related personal health technologies however, the implementation of regression modelling on users' own devices may better mitigate the potential risks entailed in large scale data sharing while maintaining the possibility of providing personalized real-time feedback. Indeed, this is the approach we are currently following in developing the next iteration of MUBS and would recommend as best practice for any larger-scale deployment of similar mHealth technologies.

## 6 CONCLUSION

This article has reported clinical outcome and usage patterns from an 8-week clinical feasibility study involving 29 patients diagnosed with a depressive disorder. In line with other studies of smartphone technologies for mental health [18], we observed a statistically significant reduction in symptom severity during the study. We performed a sub-group analysis comparing the patients who experienced improved depressive symptoms (*beneficiaries*) from those who displayed minimal or no improvement (*non-beneficiaries*). This analysis helped understand differences in system usage between the two groups. There were two main differences between the two groups. First, we saw a more substantial degree of *personalization* in the *beneficiaries*. The *beneficiaries* reduced their overall use of activities from the standard catalog and instead created and scheduled their own individual activities. These activities were mainly within the “daily living” category, such as medication reminders, sleep hygiene activities, or healthy eating. Second, the *beneficiaries* displayed a more diverse activation by scheduling and performing a more balanced and broad set of activities across different types. They took inspiration to plan activities within life areas, which seems to be neglected.

This study confirms the findings of other studies that a smartphone-based approach to BA might be a feasible approach for supporting mental health problems, especially depression. Further clinical research with a randomized control study setup is needed to provide clinical evidence. Based on our detailed analysis of what seems to benefit the *beneficiaries*, we recommend that smartphone-based systems for BA incorporate support for personalization, allowing patients to customize their activation while still having access to the catalog of standardized and generic activities for inspiration. As such, recommender technology may play a central role in helping the patient create a more diversified schedule by creating awareness of different activities and supporting the planning of activities within less explored areas.

## REFERENCES

- [1] Saeed Abdullah, Elizabeth L. Murnane, Mark Matthews, and Tanzeem Choudhury. 2017. *Circadian Computing: Sensing, Modeling, and Maintaining Biological Rhythms*. Springer International Publishing, Cham, 35–58. [https://doi.org/10.1007/978-3-319-51394-2\\_3](https://doi.org/10.1007/978-3-319-51394-2_3)
- [2] American Psychiatric Association et al. 2013. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. American Psychiatric Publishing, Washington, DC.
- [3] Sepideh Bakht, Tahereh Mahdavi Haji, Ensiyeh Ghasemian Shirvan, and Hamed Ekhtiari. 2015. The persian checklist of pleasant events (PCPE): Development, validity and reliability. *Iran. J. Psychiatr.* 10, 4 (2015), 246–264.
- [4] Jakob E. Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. 2013. 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)*. Association for Computing Machinery, New York, NY, 2627–2636. <https://doi.org/10.1145/2470654.2481364>
- [5] Aaron T. Beck. 1979. *Cognitive Therapy of Depression*. Guilford Press.
- [6] Michelle Nicole Burns, Mark Begale, Jennifer Duffecy, Darren Gergle, Chris J. Karr, Emily Giangrande, and David C. Mohr. 2011. Harnessing context sensing to develop a mobile intervention for depression. *J. Med. Internet Res.* 13, 3 (12 Aug 2011), e55. <https://doi.org/10.2196/jmir.1838>

- [7] Christopher Burton, Brian McKinstry, Aurora Szentagotai Tatar, Antoni Serrano-Blanco, Claudia Pagliari, and Maria Wolters. 2013. Activity monitoring in patients with depression: A systematic review. *J. Affect. Disord.* 145, 1 (2013), 21–28. <https://doi.org/10.1016/j.jad.2012.07.001>
- [8] Peter A. Creed and Sean R. Macintyre. 2001. The relative effects of deprivation of the latent and manifest benefits of employment on the well-being of unemployed people. *J. Occupat. Health Psychol.* 6, 4 (2001), 324. <https://doi.org/10.1037/1076-8998.6.4.324>
- [9] Pim Cuijpers, Annemieke van Straten, and Lisanne Warmerdam. 2007. Behavioral activation treatments of depression: A meta-analysis. *Clin. Psychol. Rev.* 27, 3 (2007), 318–326. <https://doi.org/10.1016/j.cpr.2006.11.001>
- [10] Jennifer Dahne, Anahi Collado, C. W. Lejuez, Cristina M. Risco, Vanessa A. Diaz, Lisa Coles, Jacob Kustanowitz, Michael J. Zvolensky, and Matthew J. Carpenter. 2019. Pilot randomized controlled trial of a spanish-language behavioral activation mobile app Aptivate for the treatment of depressive symptoms among united states Latinx adults with limited English proficiency. *J. Affect. Disord.* 250, (Feb. 2019), 210–217. <https://doi.org/10.1016/j.jad.2019.03.009>
- [11] Jennifer Dahne, C. W. Lejuez, Vanessa A. Diaz, Marty S. Player, Jacob Kustanowitz, Julia W. Felton, and Matthew J. Carpenter. 2019. Pilot randomized trial of a self-help behavioral activation mobile app for utilization in primary care. *Behav. Ther.* 50, 4 (2019), 817–827. <http://doi.org/10.1016/j.beth.2018.12.003>
- [12] Mark Deady, David Johnston, David Milne, Nick Glozier, Dorian Peters, Rafael Calvo, and Samuel Harvey. 2018. Preliminary effectiveness of a smartphone app to reduce depressive symptoms in the workplace: Feasibility and acceptability study. *JMIR Mhealth Uhealth* 6, 12 (04 Dec. 2018), e11661. <https://doi.org/10.2196/11661>
- [13] Anind K. Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin-Hyuk Hong, and Julian Ramos. 2011. 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)*. Association for Computing Machinery, New York, NY, 163–172. <https://doi.org/10.1145/2030112.2030135>
- [14] Amandeep Dhir, Yossiri Yossatorn, Puneet Kaur, and Sufen Chen. 2018. Online social media fatigue and psychological wellbeing—A study of compulsive use, fear of missing out, fatigue, anxiety and depression. *Int. J. Inf. Manage.* 40 (2018), 141–152. <http://doi.org/10.1016/j.ijinfomgt.2018.01.012>
- [15] Sona Dimidjian, Manuel Barrera, Christopher Martell, Ricardo F. Muñoz, and Peter M. Lewinsohn. 2011. The origins and current status of behavioral activation treatments for depression. *Annu. Rev. Clin. Psychol.* 7, 1 (2011), 1–38. <https://doi.org/10.1146/annurev-clinpsy-032210-104535>
- [16] Michael H. Ebert, Peter T. Loosen, Barry Nurcombe, and J. F. Leckman. 2000. *Current Diagnosis & Treatment in Psychiatry*. Lange Medical Books/McGraw–Hill.
- [17] Amanda Edwards-Stewart. 2012. Using technology to enhance empirically supported psychological treatments: Positive activity jackpot. *Arch. Med. Psychol.* 3, 2 (2012), 60–66.
- [18] Joseph Firth, John Torous, Jennifer Nicholas, Rebekah Carney, Abhishek Pratap, Simon Rosenbaum, and Jerome Sarris. 2017. The efficacy of smartphone-based mental health interventions for depressive symptoms: A meta-analysis of randomized controlled trials. *World Psychiatr.* 16, 3 (2017), 287–298. <http://doi.org/10.1002/wps.20472>
- [19] Theresa Fleming, Lynda Bavin, Mathijs Lucassen, Karolina Stasiak, Sarah Hopkins, and Sally Merry. 2018. Beyond the trial: Systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J. Med. Internet Res.* 20, 6 (06 Jun. 2018), e199. <https://doi.org/10.2196/jmir.9275>
- [20] 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. 2018. A mobile app-based intervention for depression: End-user and expert usability testing study. *J. Med. Internet Res.* 20, 8 (2018), 1–12. <https://doi.org/10.2196/mental.9445>
- [21] Toshi A. Furukawa, Masaru Horikoshi, Hirokazu Fujita, Naohisa Tsujino, Ran Jinnin, Yuki Kako, Sei Ogawa, Hirotohi Sato, Nobuki Kitagawa, Yoshihiro Shinagawa, Yoshio Ikeda, Hissei Imai, Aran Tajika, Yusuke Ogawa, Tatsuo Akechi, Mitsuhiko Yamada, Shinji Shimodera, Norio Watanabe, Masatoshi Inagaki, and Akio Hasegawa. 2018. Cognitive and behavioral skills exercises completed by patients with major depression during smartphone cognitive behavioral therapy: Secondary analysis of a randomized controlled trial. *JMIR Ment. Health* 5, 1 (11 Jan. 2018), e4. <https://doi.org/10.2196/mental.9092>
- [22] Toshi A. Furukawa, Hissei Imai, Masaru Horikoshi, Shinji Shimodera, Takahiro Hiroe, Tadashi Funayama, and Tatsuo Akechi. 2018. Behavioral activation: Is it the expectation or achievement, of mastery or pleasure that contributes to improvement in depression? *J. Affect. Disord.* 238, (May 2018), 336–341. <https://doi.org/10.1016/j.jad.2018.05.067>
- [23] Franz Gravenhorst, Amir Muaremi, Jakob Bardram, Agnes Grünerbl, Oscar Mayora, Gabriel Wurzer, Mads Frost, Venet Osmani, Bert Arnrich, Paul Lukowicz, and Others. 2015. Mobile phones as medical devices in mental disorder treatment: An overview. *Pers. Ubiq. Comput.* 19, 2 (2015), 335–353. <https://doi.org/10.1007/s00779-014-0829-5>
- [24] Allison G. Harvey. 2008. Sleep and circadian rhythms in bipolar disorder: Seeking synchrony, harmony, and regulation. *Am. J. Psychiatr.* 165, 7 (2008), 820–829. <https://doi.org/10.1176/appi.ajp.2008.08010098>
- [25] Ainslie Hatch, Julia E. Hoffman, Ruth Ross, and John P. Docherty. 2018. Expert consensus survey on digital health tools for patients with serious mental illness: Optimizing for user characteristics and user support. *J. Med. Internet Res.* 20, 6 (2018), e46. <https://doi.org/10.2196/mental.9777>

- [26] Keith Hawton and Kees van Heeringen. 2009. Suicide. *Lancet (Lond. Engl.)* 373, 9672 (Apr. 2009), 1372–81. [https://doi.org/10.1016/S0140-6736\(09\)60372-X](https://doi.org/10.1016/S0140-6736(09)60372-X)
- [27] Anna Huguet, Sanjay Rao, Patrick J. McGrath, Lori Wozney, Mike Wheaton, Jill Conrod, and Sharlene Rozario. 2016. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS ONE* 11, 5 (2016), 1–19. <https://doi.org/10.1371/journal.pone.0154248>
- [28] Neil S. Jacobson, Keith S. Dobson, Paula A. Truax, Michael E. Addis, Kelly Koerner, Jackie K. Gollan, Eric Gortner, and Stacey E. Prince. 1996. A component analysis of cognitive-behavioral treatment for depression. *J. Consult. Clin. Psychol.* 64, 2 (1996), 295. <http://doi.org/10.1037//0022-006x.64.2.295>
- [29] Jonathan W. Kanter, Rachel C. Manos, William M. Bowe, David E. Baruch, Andrew M. Busch, and Laura C. Rusch. 2010. What is behavioral activation?: A review of the empirical literature. *Clin. Psychol. Rev.* 30, 6 (2010), 608–620. <http://doi.org/10.1016/j.cpr.2010.04.001>
- [30] Patricia P. Katz and Edward H. Yelin. 2001. Activity loss and the onset of depressive symptoms: Do some activities matter more than others? *Arthrit. Rheumat.* 44, 5 (2001), 1194–1202. [https://doi.org/10.1002/1529-0131\(200105\)44:5<1194::AID-ANR203>3.0.CO;2-6](https://doi.org/10.1002/1529-0131(200105)44:5<1194::AID-ANR203>3.0.CO;2-6)
- [31] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. 2011. 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)*. ACM, New York, NY, 3063–3072. <https://doi.org/10.1145/1978942.1979396>
- [32] Leon Kreitzman and Russell Foster. 2011. *The Rhythms of Life: The Biological Clocks that Control the Daily Lives of Every Living Thing*. Profile Books.
- [33] Soomi Lee, Rachel E. Koffer, Briana N. Sprague, Susan T. Charles, Nilam Ram, and David M. Almeida. 2018. Activity diversity and its associations with psychological well-being across adulthood. *J. Gerontol. Ser. B, Psychol. Sci. Soc. Sci.* 73, 6 (2018), 985–995. <http://doi.org/10.1093/geronb/gbw118>
- [34] C. W. Lejuez, Derek R. Hopko, James P. LePage, Sandra D. Hopko, and Daniel W. McNeil. 2001. A brief behavioral activation treatment for depression. *Cogn. Behav. Pract.* 8, 2 (Mar. 2001), 164–175. [https://doi.org/10.1016/S1077-7229\(01\)80022-5](https://doi.org/10.1016/S1077-7229(01)80022-5)
- [35] Peter M. Lewinsohn. 1974. *A Behavioral Approach to Depression*. New York University Press, New York, NY, 150–172 pages.
- [36] Peter M. Lewinsohn and Michael Graf. 1973. Pleasant activities and depression. *J. Consult. Clin. Psychol.* 41, 2 (1973), 261. <https://doi.org/10.1037/h0035142>
- [37] Kien Hoa Ly, Elsa Janni, Richard Wrede, Mina Sedem, Tara Donker, Per Carlbring, and Gerhard Andersson. 2015. Experiences of a guided smartphone-based behavioral activation therapy for depression: A qualitative study. *Internet Intervent.* 2, 1 (2015), 60–68. <https://doi.org/10.1016/j.invent.2014.12.002>
- [38] Kien Hoa Ly, Naira Topooco, Hanna Cederlund, Anna Wallin, Jan Bergstrom, Olof Molander, Per Carlbring, and Gerhard Andersson. 2015. Smartphone-supported versus full behavioural activation for depression: A randomised controlled trial. *PLoS ONE* 10, 5 (2015), 1–16. <https://doi.org/10.1371/journal.pone.0126559>
- [39] Kien Hoa Ly, Anna Trüschel, Linnea Jarl, Susanna Magnusson, Tove Windahl, Robert Johansson, and Per Carlbring. 2014. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: A randomised controlled trial. *Br. Med. J.* 4 (2014), e003440. <https://doi.org/10.1136/bmjopen-2013-003440>
- [40] D. J. MacPhillamy and P. M. Lewinsohn. 1972. The measurement of reinforcing events. In *Proceedings of the 80th Annual Convention of the American Psychological Association*, Vol. 7. American Psychological Association, Washington, DC, 399–400.
- [41] Douglas J. MacPhillamy and Peter M. Lewinsohn. 1982. The pleasant events schedule: Studies on reliability, validity, and scale intercorrelation. *J. Consult. Clin. Psychol.* 50, 3 (1982), 363–380. <https://doi.org/10.1037/0022-006X.50.3.363>
- [42] Akio Mantani, Tadashi Kato, Toshi A. Furukawa, Masaru Horikoshi, Hissei Imai, Takahiro Hiroe, Bun Chino, Tadashi Funayama, Naohiro Yonemoto, Qi Zhou, and Nao Kawanishi. 2017. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: Randomized controlled trial. *J. Med. Internet Res.* 19, 11 (3 Nov 2017), e373. <https://doi.org/10.2196/jmir.8602>
- [43] Gabriela Marcu, Jakob E. Bardram, and Silvia Gabrielli. 2011. A framework for overcoming challenges in designing persuasive monitoring and feedback systems for mental illness. In *Proceedings of the 2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'11)*. IEEE, 1–8. <http://doi.org/10.4108/icst.pervasivehealth.2011.246097>
- [44] Christopher R. Martell, Sonja Dimidjian, and Ruth Herman-Dunn. 2013. *Behavioral Activation for Depression: A Clinician's Guide*. Guilford Press.
- [45] Charlotte M. McKercher, Michael D. Schmidt, Kristy A. Sanderson, George C. Patton, Terence Dwyer, and Alison J. Venn. 2009. Physical activity and depression in young adults. *Am. J. Prevent. Med.* 36, 2 (2009), 161–164. <https://doi.org/10.1016/j.amepre.2008.09.036>
- [46] Merete M. Mørch and Nicole K. Rosenberg. 2005. *Kognitiv Terapi: Modeller og Metoder*. Gyldendal A/S.
- [47] Ehimwenma Nosakhare and Rosalind Picard. 2020. Toward assessing and recommending combinations of behaviors for improving health and well-being. *ACM Trans. Comput. Healthcare* 1, 1, Article 4 (Mar. 2020), 29 pages. <https://doi.org/10.1145/3368958>
- [48] Francisco Nunes. 2019. From medicalized to mundane self-care technologies. *Interactions* 26, 3 (Apr. 2019), 67–69. <https://doi.org/10.1145/3319374>

- [49] Mashfiqui Rabbi, Min S. H. Aung, Geri Gay, M. Cary Reid, and Tanzeem Choudhury. 2018. Feasibility and acceptability of mobile phone-based auto-personalized physical activity recommendations for chronic pain self-management: Pilot study on adults. *J. Med. Internet Res.* 20, 10 (26 Oct. 2018), e10147. <https://doi.org/10.2196/10147>
- [50] David A. Richards, David Ekers, Dean McMillan, Rod S. Taylor, Sarah Byford, Fiona C. Warren, Barbara Barrett, Paul A. Farrand, Simon Gilbody, Willem Kuyken, Heather O'Mahen, Ed R. Watkins, Kim A. Wright, Steven D. Hollon, Nigel Reed, Shelley Rhodes, Emily Fletcher, and Katie Finning. 2016. Cost and outcome of behavioural activation versus cognitive behavioural therapy for depression (COBRA): A randomised, controlled, non-inferiority trial. *The Lancet* 388, 10047 (Aug. 2016), 871–880. [https://doi.org/10.1016/S0140-6736\(16\)31140-0](https://doi.org/10.1016/S0140-6736(16)31140-0)
- [51] Kenneth L. Rider, Larry W. Thompson, and Dolores Gallagher-Thompson. 2016. California older persons pleasant events scale: A tool to help older adults increase positive experiences. *Clin. Gerontol.* 39, 1 (2016), 64–83. <https://doi.org/10.1080/07317115.2015.1101635>
- [52] Darius A. Rohani, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E. Bardram. 2018. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review. *JMIR Mhealth Uhealth* 6, 8 (13 Aug. 2018), e165. <https://doi.org/10.2196/mhealth.9691>
- [53] Darius A. Rohani, Andrea Quemada Lopategui, Nanna Tuxen, Maria Faurholt-Jepsen, Lars V. Kessing, and Jakob E. Bardram. 2020. MUBS: A personalized recommender system for behavioral activation in mental health. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20)*. Association for Computing Machinery, New York, NY, 1–13. <https://doi.org/10.1145/3313831.3376879>
- [54] D. A. Rohani, A. Springer, V. Hollis, J. E. Bardram, and S. Whittaker. 2020. Recommending activities for mental health and well-being: Insights from two user studies. *IEEE Trans. Emerg. Top. Comput.* (2020), <https://doi.org/10.1109/tetc.2020.2972007>
- [55] Darius A. Rohani, Nanna Tuxen, Andrea Quemada Lopategui, Maria Faurholt-Jepsen, Lars V. Kessing, and Jakob E. Bardram. 2019. 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)*. ACM, New York, NY, 282–291. <https://doi.org/10.1145/3329189.3329214>
- [56] Darius Adam Rohani, Nanna Tuxen, Andrea Quemada Lopategui, Lars Vedel Kessing, and Jakob Eyvind Bardram. 2018. Data-driven learning in high-resolution activity sampling from patients with bipolar depression: Mixed-methods study. *JMIR Ment. Health* 5, 2 (2018), e10122. <https://doi.org/10.2196/10122>
- [57] John Rooksby, Alistair Morrison, and Dave Murray-Rust. 2019. 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 (CHI'19)*. Association for Computing Machinery, New York, NY, 1–14. <https://doi.org/10.1145/3290605.3300655>
- [58] Mattias Rost, John Rooksby, Alexandra Weilenmann, Thomas Hillman, Pål Dobrin, and Juan Ye. 2016. Mobile wellbeing. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI'16)*. Association for Computing Machinery, New York, NY, Article 137, 3 pages. <https://doi.org/10.1145/2971485.2987676>
- [59] Sohrab Saeb, Mi Zhang, Christopher J. Karr, Stephen M. Schueller, Marya E. Corden, Konrad P. Kording, and David C. Mohr. 2015. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *J. Med. Internet Res.* 17, 7 (2015), 1–11. <https://doi.org/10.2196/jmir.4273>
- [60] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, and Corina Sas. 2019. 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)*. ACM, New York, NY, Article 245, 17 pages. <https://doi.org/10.1145/3290605.3300475>
- [61] Ziggi Ivan Santini, Ai Koyanagi, Stefanos Tyrovolas, Catherine Mason, and Josep Maria Haro. 2015. The association between social relationships and depression: A systematic review. *J. Affect. Disord.* 175 (2015), 53–65. <http://doi.org/10.1016/j.jad.2014.12.049>
- [62] Tobias Sonne, Jörg Müller, Paul Marshall, Carsten Obel, and Kaj Grønbaek. 2016. Changing family practices with assistive technology: MOBERO improves morning and bedtime routines for children with ADHD. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'16)*. Association for Computing Machinery, New York, NY, 152–164. <https://doi.org/10.1145/2858036.2858157>
- [63] Aaron Springer, Victoria Hollis, and Steve Whittaker. 2018. Mood modeling: Accuracy depends on active logging and reflection. *Pers. Ubiq. Comput.* 22, 4 (2018), 723–737. <https://doi.org/10.1007/s00779-018-1123-8>
- [64] Linda Teri and Rebecca G. Logsdon. 1991. Identifying pleasant activities for alzheimer's disease patients: The pleasant events schedule-AD. *The Gerontol.* 31, 1 (1991), 124–127. <http://doi.org/10.1093/geront/31.1.124>
- [65] Megan Teychenne, Kylie Ball, and Jo Salmon. 2010. Sedentary behavior and depression among adults: A review. *Int. J. Behav. Med.* 17, 4 (2010), 246–254. <https://doi.org/10.1007/s12529-010-9075-z>
- [66] John Torous, Rohn Friedman, and Matcheri Keshavan. 2014. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR mHealth uHealth* 2, 1 (21 Jan. 2014), e2. <https://doi.org/10.2196/mhealth.2994>
- [67] John Torous and Laura Weiss Roberts. 2017. Needed innovation in digital health and smartphone applications for mental health: Transparency and trust. *JAMA Psychiatry* 74, 5 (2017), 437–438. <http://doi.org/10.1001/jamapsychiatry.2017.0262>
- [68] John Torous, Jessica Woodyatt, Matcheri Keshavan, and Laura M. Tully. 2019. A new hope for early psychosis care: The evolving landscape of digital care tools. *Br. J. Psychiatr.* 214, 5 (2019), 269–272. <https://doi.org/10.1192/bjp.2019.8>

- [69] Yonatan Vaizman, Katherine Ellis, Gert Lanckriet, and Nadir Weibel. 2018. ExtraSensory app: Data collection in-the-wild with rich user interface to self-report behavior. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18)*. Association for Computing Machinery, New York, NY, 1–12. <https://doi.org/10.1145/3173574.3174128>
- [70] Fabian Wahle, Tobias Kowatsch, Elgar Fleisch, Michael Rufer, and Steffi Weidt. 2016. Mobile sensing and support for people with depression: A pilot trial in the wild. *JMIR mHealth uHealth* 4, 3 (2016), e111. <https://doi.org/10.2196/mhealth.5960>

Received July 2020; revised February 2021; accepted April 2021